

CCAGGTCCAAC TGCACTCGGTTCTATCGATTGAAT TCCCCGGGGATCCTCTAGAGATCCCT
 CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCCTGACGGGTGCAAGGGAGGCTCCTGTGGGA
 CAGGCCAGGCAGGTGGGGCTCAGGAGGTGCCTCCAGCGCGCCAGTGGGCCTGAGGCCCGACG
 AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCTTGGGC
 TCCAGCAGCATCAGCAGCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
 AGCAGCTCCTGCCCCGTGTCGGGGGGAAGTACTGATTCTCCTCCGCAAGGCCACCCAGAGGAGA
 AGGCCACCCCGCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
 TCTGGTGTTGGCAGTGGGCGGCACAGCAGCCTACCGCCCGCGCGGTTAGGGTTGTGTGCT
 TCCCGGCTCAGCGGGACCCTGTCTCCGAGTCTGTCGTGCAGCGTGTGTACCAGCCCTTCC
 TCACCAGCTCGCAGCGGGCACCGGGCCCTGCAGCACCCTACCGAACCATTATATAGGACCGCCTAC
 CGCCGACGCTTGGGCTGGCCCCCTGCCAGGCCCTCGCTACGCGTGTGTCGCCCGGCTGGAAGAG
 GACCAGCGGGCTTCTGGGGCTGTGGAGCAGCAATATGCCAGCCGCATGCCGGAACGGAG
 GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCACG
 TCAGATGTGGATGAATGCAAGTGTAGGAGGGGGCGGCTGTCCCCAGCGTGTGCATCAACACCGC
 CGGCAGTTACTGTGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
 TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCATGAAG
 GAAGAAGTGCAGAGCTGCAGTCCAGGGTGGACCTGTGGAGGAGAAGTGCAGCTGGCTGTCT
 GGCCCACTCAGACGCCCTGGCTCGCAGGCATGGAGCATGGGCTCCCGGACCCCGGACGC
 TCTTGGTGCACTCCTTCCAGCAGCTCGGCCGCACTGCACTCCCTGAGCGAGCAGATTTCCTTC
 CTGGAGGAGCAGCTGGGGTCTGCTCTCTGAAGAAAGACTCGTGACTGCCACGCGCCCCAGG
 CTGGACTGAGCCCCCTACGCGCGCCCTGCAGCCCCCATGCCCCCTGCCAACATGTGGGGGTC
 CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCCTTTTCTCCTC
 CCTTCTCCTCGGGAGGGTCCCCAGACCCTGGCATGGATGGGCTGGGATTTTTTTTGTGAAT
 CCACCCCTGGGTACCCCCACCTTGGTTACCCCAACGGCATCCCAAGGCAGGTGGGCCCTCA
 GCTGAGGGAAGGTACGAGTTCCCTTGTGTGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
 CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGTGTCCTGACCCCCAGCACAAATAAAATGAAA
 CTGTGAAAGAAAAA
 AGGGCGCGCGCAGCTCT
 AGAGTCAGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTATTATTCAGCTTATAATGCT
 TACAATAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVSGRHRARLPARPLGCVLSRAHGDPV
SESFVQVRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCPVKGPPRPVAPNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTCGCCCTCGCCTCCCCGCAGAGTCCCTCGCGGCAGAGATGTGTGTGGG
 GTCAGCCACGGCGGGGACTATGGTGAAATTCGGCGCTCAGCAGTACTGGCCCTGATC
 CGGTTCTTGGTGCCTTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCCTTGA
 CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGTACGGGCTGGCGT
 ACTCCCTCATGAAGTTCTTACGGGTCCCATGAGTGACTTCAAAAATGTGGGCTGTGTGTT
 GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCTGTGTATGGTGGTGGCAGGGGCCATCGC
 TGCCGTCTTACACACTGTGATAGCTTATAGTGATTTAGGATACATATTATCAATAAACTGC
 ACCATGTGGACGAGTCGGTGGGGAGCAAGAAGAGAAGGGCCTCTCTGTACCTCGCCGCCCTT
 CCTTTCATGGACCAATGGCATGGACCCATGCTGGCATTCTCTTAAACCAAAATACAGTTT
 CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTGTAGCCATTTTGC
 TTCAGAGTCACCTGGAATGCGGGAGCCCTGCTCATCCCGATCTCTCTCTGTACATGGGC
 GCACTTGTGCGCTGCACCAACCTGTGCTTGGGCTACTACAAGAACATTACGACATCATCCC
 TGACAGAAGTGGCCCGGAGCTGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGTT
 GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
 TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGCAGTGGCGATTTTGACAGCCACATA
 CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGTGTGTATCTCTGTTCTTG
 ACAAGAATAAACCCAGCAACAAACTGGTGAGCACGAGCAACAGTCAGCGGCAGCCACATC
 AAGAAGTTACCTTCGTCTGCATGGCTCTGTCACTCAGCTCTGTTCGTGTATGTTTTGGAC
 ACCCAAGCTGTCTGAGAAAATCTTGATAGACATCATCGAGTGGACTTTGCTTTTGAGAACT
 TCTGTGTTCTTCTTTGGGATCTCTCTCTCTTCCAGTTCCAGTCAGAGTGGGCGCAT
 CTCACCGGTTGGCTGATGACACTGAAGAAAACCTTCGTCTTGGCCCCAGCTCTGTCTGG
 GATCATCGTCTCTACGCCACCTCGTGGTCTACCTACCTCGGGGTGCACGGTGCAGCCC
 TGGGCTGGGCTCCCTCTGGCGGGCTTTGTGGGAGAATCACCATGGTCGCCATCGCTGGG
 TGCTATGTCTACCGGAAGCAGAAAAGAAGATGGAGAATGAGTCGGCCACGAGGGGGGAAGA
 CTCCTGCCATGACGACATGCTCCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAG
 AGAATGAATAAGGCACGGGACGCCATGGGCATGACAGGACGGTCAGTCAGGATGACACTTC
 GGCATCATCTCTCCCTCTCCCATCGTATTTTGTTCCTTTTGTGTTTTGTTTTGTTAAAT
 GAAAAGAGGCTTGATTTAAAGGTTTCGTGTCAATTCTTAGCATCTCGGATGCTCACACT
 GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
 TATCATCCCTGCTCACGAAAACCCAAAAGACACAGCTGCCTCAGCGTTGACGTTGTGTCC
 TCCTCCCCGGACAATCTCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCTCGGT
 CACCTGACAGCAGGCGCACAGACTCTCTGTCCCTTTCATCGCTCTTAAGAATCAACAGG
 TTAATAACTCGGCTTCTTTGATTTGCTTCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
 CGGTGGCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
 GAGGCGGGTGGCAGCTGCAGCCCGAGTCCCGTTCACACTGAGGAACGGAGACCTGTGAC
 CACAGCAGGCTGACAGATGGACAGAATCTCCGTAGAAAGGTTGGTTTGAATGCCCGGG
 GGCAGCAAACTCAGCTGTTGAATGATAGCATTTCACTCTGCGTCTCTCTAGATCTGAGCAA
 GGTGCAGTTCTCACCCACCGTGTATATACATGAGCTAACTTTTAAATTTGTCAAAAA
 GCGCATCTCCAGATTCAGACCCCTGCCGATGACTTTTCTGAAAGCTTGCTTTTCCCTCGC
 CTTTCTGAAAGTTCGATTTAGAGCGAGTCACATGGAGCATCTTAACTTTGCAATTTAGTTTT
 TACAGTAACCTGAAGCTTTAAGTCTCATCCAGCATCTAATGCCAGTTGCTGTAGGGTAA
 TTTTGAAGTAGATATATACCTGGTTCTGCTATCCCTTAGTCATAACTCTGCGGTACAGGTAA
 TTGAGAATGTACTACGGTACTTCCCTCCACACCATACGATAAAGCAAGACATTTATAACG
 ATACAGAGTCACTATGTGGTCTCTCCGAAATAACGCATTCGAAATCCATGCAAGTCAGTA
 TATTTTCTAAGTTTGTGAAAGCAGGTTTTTCTTTAAAAAAATATAGACACGGTTCACT
 AAATTTGATTTAGTCAGAACTTCTAGACTGAAAGAACCTAAACAAAAAATATTTTAAAGATA
 TAAATATATGCTGTATATGTTATGTAAATTTATTTTAGGCTATAATACATTTCTATTTTCG
 ATTTTCAATAAAATGCTCTAATACAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSEDFKNVGLVFVNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSIDLGYIINKLHHVDESV
GSKTRRAFLYLAAFPMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCTTLCLGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPIVNLVFSRDLGGSSAATEAVAILTATYPVGHMPYGLWLTEIRAVYPAFDKNPNPSN
KLVSTSNVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFaelcVVPLR
IFSFFPVPVTRAHLTGWMLTLKKTfVLAPSSVLRIIVLIASLVLPYLGvHGATLGVGSLl
AGFVGESTMVAIAACYVYRKQKKKMenESATEGEDSAMTDMPTTEEVTDIVEMRENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGAAATCCGTGCTGTGTATCCTGCTTTG
ACAAGAATAACCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNAAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGTGGCTGATGACACTGAAGAAAACCTTCGTC

FIGURE 6

TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGGAAAACAGTNACGGGCAGCCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTGAGAAAAATTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTTCCTTTCGGGATTTTCT
CCTTTTTCCCAGTTCAGTTCACAGNAGGGCGCATCTACCGGGNGGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCCTACCTGGGGGTGCACGGTGCGACCCTGGGCGTGGGTTCCTCCTGGCGGGCA

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 7

TATTTCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGAACTGAAGAAA
ACCTTNGTCCTTGCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT
CCTACCCCTACCTGGGGGTGCACGGTGAGAC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 8

GCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCGCCAGCCGCAACTCCACCGTGAGCCGCTCATCTTACAGTTCTTCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGG
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT
TCTTTTTCACCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG
AATGGGTTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCCTACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTCACCAACATCTGGTCTACTTCGCGCTCGTGGGCTCCTTCTCTTCA
TCCTCATCCAGCTGGTGTGCTCATCGACITTGCGCACTCCTGGAACAGCGGTGGTGGGC
AAGGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCAGGCCCTCTTCTTCTTCACTCTCCTCTT
CTACTTGCTGTCGATCGCGGCCGTGGCGCTGATGTTATGTACTACACTGAGCCAGCGGCT
GCCACGAGGGCAAGGTCTTATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCTGCCAAGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTGAT
CACCTCTACACCATGTTTGTCACTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCATTTGCCAACCCAGCTGGGCAACGAGACAGTTTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCTCCTGTGCACCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGGCTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGT
GGCCTCACTGCACGTGATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCGGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCTTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTGAAGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGTGACAGCCAACCT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCCTTTCAAGGTCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCTTCCCCCTCCTCCTGT
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAATCCCAACACAGTGGGGCATCCGGCACTGAAGCCCTGGTGT
CCTGGTCAGTCCCCAGGGGACCTGCCCCCTTCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTCTAATAAACAGCCAGTGCCTGTAAAAAAA

FIGURE 9

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGLVLSIIMLSPGVE
SGLYKLPWVCEEAGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRD
PRAAIQNGWFFKFLILVGLTVGAFYIPDGSFTNIWIFYFGVVGSLFLILIQLVLLIDFAHSW
NQRWLGKAEECDRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFC
VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVA
ACEGRAFDNEQDGVITYSYFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICAS
WAGLLLYLWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

GAGCGAGGCGCGGGACATGAAGGTGTGGGTGTCGAGCCCTCTGGCAGAGGGTTAACTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGGGCAGCTCCGCGAGGACTTGA
AGTCTTGAGCGCTCAAGTTTGTCCGTAGGTGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCCGAGTCTTTGTCTGCCGAAGCTG
TGACTGCCGATTGGAAGTCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGGAAATACAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCCA
GAAATTTTATCATACCGGCTTTGATGCTGTGCAATCTGCACATCGTGTGCCACAGGAGGCTT
CATTGTTTATGGCTGCGGCTGGGGTTGGAGAACTGCAGTGTGTTTGATCATATTAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAGATGCCTTAAGCCATTTTGTAA'TTGCAGGA
GCTGTACCGGAAGTCTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAAAGAAGGATCGAAAGGCATCCATGAGCTAAACTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAACAAAGACAAGGACTGAAAGTGCTCTGAACCTTGAACCTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTCTTTTCTT
TTTAAGTAAGAATTGGGCTGTTGTACTCTCATCTTACTTATCTTAAATCTTAAATTAACATACT
TATGTTTGTATTAATCTATCATATATGATCATACATGATATCAACCCACTAGATTTTAA
GCAGTAAATAAAACATTTCGCAAAAGATTAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318

><subunit 1 of 1, 285 aa, 1 stop

><MW: 32190, pI: 9.03, NX(S/T): 2

MEVPPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAETIYHNRFDVQSAH
RAATRGFIIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDLSHFVIAGAVTGSLSFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRQLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTTCGTTTCATG
GCTGGCGCCGAACC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
220

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCTCGAGTTTTTGTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 14

GAGCCGCCGCCGCCGCCGCCGCCGCCGCTGCAGCCCCAGGCCGCCGCCGCCACCGTCT
GCGTTGCTGCCCCGCCCTGGGCCAGGCCCAAAGCAAGGACAAAGCAGCTGTGAGGAACCT
CCGCCGAGTTCGAATTTACGTGCAGCTGCCGCCAACACAGGTTCCAAGATGGTTTTGCGGGG
GCTTCGCGTGTTCGAAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGTTAGTCTG
CTGCTAAATTGGAATTGTCGCTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGGTCATTGCAGTGGGCATCTTCTTGTCTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAAACATCATCAGGTGTTGCTATTTTTTTATATGATTATTCTGTTACTTGTATTTATT
GTTACGTTTTCTGTATCTTGCCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAACTGCT
GTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCGCGCGAATCCTAGTGCAATCTCTTGGATGAGAAAACAAGGAAGAT
TTCCTTTCGTATATGATCTTGTTCACCTTCTGTAATTTTCTGTTAAGCTCCATTGCCAGT
TTAAGGAAGGAAACATATCTGGAAGTACCTTATTGATAGTGGAAATATATATTTTTACT
CTATGTTTCTCTACATGTTTTTTCTTCCGTTGCTGAAAAATATTTGAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCATTGTGCGGCAGTGTCCACTGTGGCCTT
TCTTAGCATTTTACCTGCAGAAAACTTTGTATGGTACCAGTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCCTAC
TGGA AAAAGAGTGGAATTTATTA AAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCCAATTTTTTTTGGTCTTTTATGAAAGATTGTTGTTGTTAAAAAGTGTAGTA
TAAAAATGATAATTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTTAGGAAATGTGGTTTAAATTTTTGACTTTTACAGGTAAGTGCAAGGAGAAGTGGTTT
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA
AAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT
GTCAAACTAAGCATATTTGAATATGATCTCCATAATTTGAAATGAAATCGTATTGTGTG
GCTCTGTATATTCTGTTAAAAAATTAAGGACAGAAACCTTCTTGTGTATGCATGTTTGA
ATTAAAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVIAGVIGFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFRVGGIGLFFSFTEILGVWL
TYRYRNQKDPRANPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTATATGATTATTCGTAANT
TGTATTTATTGTTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAAT
NTAAACTGCTGTGGGTTCCGAAGTGTTAACCCTAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2

FIGURE 17

AATCCCAAATCCCCAATTTTTTTGGNCTTTTtagggAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTGAAG
TCTAAAAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAAGTTCTCAGTATTGTAAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCCATAAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181

ATGATTATTCTGTTACTGTATTATTGTTTCAGTTTATGGTATCTTGCCTTGTTTAGCCC
CTGA AACAGGAGCAACAGGGNNCAGCTTCTCGAGGTTGGTTGGCAACAATCAGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAGTGACCACTCGTCTCGCCATGTGCTCCA
TCATAGGAGAAATATGC

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGVLWVAQMLLAASFETLQCEGPVCTEESCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRCAWQDWPLTQVTIFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ
VWKQSPQLEIRVQGASSSAAPPTLNPAPQKSAAPGTAPPEAPGPLPPPTPSSDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCACGCGTCCGCCACGGGTCCGCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGCATGGG
ATCTTACTGGGCTGCTACTCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAACAGGACCTTGGAAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTGCAAGGCTACACCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAGTACCAGGGCCG
CCTGCATGTGAGGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCTGGAGATGG
ATGACCGGAGCCACTACAGTGTGAAGTCACCTGGCAGACTCTGTATGGCAACCAAGTCTGT
AGAGATAAGATTACTAGCTTCCGTGTCCAGAACTCTGTCTCCAAGCCACAGTGACAAAC
TGGCAGCGGTTATGGCTTACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACAGGAACCCATC
AAAGTAGCAACCCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCCCTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTGCGGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAGAAAGTCTCTCCTCTTCCATTTTTGACCCCGT
CCCTGCCCTCAATTTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCA CAGACCC
AATCCTAAGGCCGAGGCCCTT CAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTTGGCCCTCTGAAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACAGGACCCACAGCCCCAAGTCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGGCCAAAGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGTATGAGCCAACCTTCCAGAATCTGGGCAACAACACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCCGCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTTCTGGCCACTGAGGGCAAAAGTGTCTGTTA AAAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCTTGCCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCTGGATAGCCCAAAGTGTCCGCCCTACCAACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACAGCTATATCCAGGATCAT
TTCTCTTCTTTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTTTCACAGGCCAGGGTTCA
GTTCTGCTCTCTCCACTATAAGTCTAATGTTCTGACTCTCTCTGGTGTCTAATAAATATCTA
ATCATAACAGC

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416

><subunit 1 of 1, 321 aa, 1 stop

><MW: 35544, pI: 8.51, NX(S/T): 0

MGILLGLLLLGHLTVDTYGRPILEVPESVTGPWKGDVNLPCYDPLQGYTQVLVKNLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRHLVSHKVPGDVSLQLSTLEMDDRSYHTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRIQLQCQARGSPPISYIWKQQTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKTEAPTMT
YPLKATSTVKQSWDWTDMGYLGETSAGPGKSLPVFAIILIIISLCCMVVFTMAYIMLCRKT
SQQEHVYEAAAR

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCAGGGGACGCGGGCGGGGCCGCTCCCGCCCGGCACAT
 GGCTGCAGCCACCTCGCGCGCACCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTCCGA
 GCGCGCCCGGCGCCCGCGGAGCCAAAGCAGCAACTGAGCGGGGAAGCGCCCGCTCCGGGGATC
 GGGATGTCCTCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACTTGGGGACTCA
 CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC
 TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTACCGATTAATGAAGGGAACCAAAA
 GTGGTGATCACTTACTCCAGTCGTATGTCTACAATAAATTGACTGAGGAACAGAAGGGCCG
 AGTGGCCTTTGCTTCCAAATTTCTGGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGC
 CCAGTGATGAGGGCGGTACACCTGTAAGGTTAAGAAATTCAGGGCTACGCTGTGGAGCCAT
 GTCATCTTAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTTGAAGGAGAGCTGAC
 AGAAGGAAGTGACTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTTGTGTATT
 ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCCTGCCCTCCCAATCTAGGATT
 GACTACAACCACCTTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
 CCAGTGCACAGCAGCAACCGAAGCTGGGAAGGAAAGCTGTGTGGTGCAGTAACCTGTACAGT
 ATGTTACAAGCATCGGCATGGTTGCAGGAGCACTGACAGGCACTAGTGCTGGCTGGAGCCCTGCTG
 ATTTTTCTCTTGCTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATAGAGGAAGAAGA
 GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTTGTGAAACCCAGCT
 CCTCTTCTCAGGCTCTCGAGCTCACGCTCTGGTTCTTCTCTCACTCGCTCCACAGCAAAAT
 AGTGCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
 CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAAGAAAGTCCACCATG
 CTAATCTGACCAAGCAGAAACCAACCCAGCATGATCCCGAGCCAGAGCAGAGCCTTCCAA
 ACGGCTCTGAATTACAATGGACTTGACTCCACGCTTTCTTAGGAGTCAAGGTCCTTTGGACTC
 TTCTCGTCATTGGAGCTCAAGTCACCGCCACACAACCCAGATGAGAGGTCATCTAAGTAGCA
 GTGAGCATTGCACGGAACAGATTAGATGAGCATTTTCTTATACAATAACCAACCAAGCAAA
 AGGATGTAAGCTGATTATCTGTGTAAGGATCTTATTTGGCCCTTAGACCCAGAGTAAGGG
 AAAGCAGGAGTCCAATCTATTGTGTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAGGTTG
 AGGTGAATATACCTAAAACITTTAATGTGGGATATTTGTATCAGTGCTTTGATTCACAATT
 TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTCTATGCATTTCTGCAAACTTATTGGATT
 ATTAGTTATTACAGACAGTCAAGCAGAACCCACAGCCTTATTACACTGTCTACACCATGTAC
 TGAGCTTAACCACTTCTAAGAACTCCAAAAAGGAAACATGTGCTTCTATTCTGACTTAAC
 TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
 AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTGTATTGAGCCCAAAATAAC
 TATGAAGAGGAGACAAAAATTTGTGACAAAGGATGTGAAGAGCTTCCATCTTCATGATGTT
 ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCTCTCAAT
 CAGATGCCTTAAGGACTTTCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
 TATCAACGCTCTTAGAAAGAATTTCTTAGAGAAAAAGGGATCAGGAATGCTGAAAGATT
 CCCACATACCAATTATAGTCTCTTCTTCTGAGAAAAATGTGAAACAGAAATTGCAAGACTGG
 GTGGACTAGAAAGGAGATTAGATCAGTTTCTTCTAATATGTCAAGGAAGGTAGCCGGGCA
 TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
 ATTGCACTCCAGCCTGGGTGACAGAGCGGACCTCCGCTCTC

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYVVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCESSTGTEPIVYVYQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 25

GTGCTTCCTTTGCTCTCTCGCGCCAGTCTCTCTCCCTGGTTCTCTCAGCCGTGTGGAGGAGAGACCCCGGA
 GACBCGGGCTGCAGTGCBCGGCGCTTCTCCCGCCTGGGGCGGCTCGCCGCTGGGCGAGTGCTGAGCGCCCTAG
 AGCCTCCCTTTCGCGCTCCCTCTCTCGCCGGCCGACAGCAGTGGGGTGTGGAGGTAGATGGGCTCCCG
 GCCCGAGGAGCGCGGTGATGTCGGCGCTGGGCAGAAAGCAGCGCCGATTCCAGCTTGCCCCGCGCGCCCGGGCG
 CCCCCTGCGAGTCCCGGTTTCAGC**CA**TGGGGACCTCTCGAGCAGCAGCACCGCCCTCGCCCTCTGCGAGCCGATC
 GCCCGCCGAGCCACAGCCACGATGATCGCGGCTCCCTTCTCTGCTTGGATTTCCTTAGCACCAACAGCAGCTGAG
 CCAGAACCAAGGGCTCGAATCTCATTTGSCACATACGCCCATTTGACCGTGCCACCGGCCAGGTGCTAACTGT
 GACAGTGTGCCAGAGGAACCTATGTCTCTGAGCATTGTACCAACACAAGCCTGCGCGTCTGCGAGCTTGCCT
 GTGCGGACCTTTACAGGATAGAGAAATGGCATAGAGAAATGCCATGACTGTAGTCAGCCATGCCCATAGCCAAATG
 ATTGAGAAATTAACCTTGTGCTGCTCTGACTGACCGAGAAATGCACTTGCCCACTGGCATGTGTCAGTCTTAACGCT
 ACCTGTGCCCCCATACGCTGTGCTCTGTGGTGTGGGTGTGCGGAAGAAAGGGACAGAGACTGAGGATGTGCGG
 TGTAAAGCAGTGTGCTCGGGTACCTTCTCAGATGTGCTTCTTAGTGTGATGAAATGCAAAAGCATACACAGACTGT
 CTGAGTCAGAACCTGGTGGTATCAAGCCGGGAGCAAGAGGAGACAGACAAAGCTCTGTGGCACATCCCGTCTCTTC
 TCCAGCTCCACCTCACCTTCCCTCGGCACAGCCATCTTTCCACGCCCTGAGCACATGGAAACCCATGAAGTCCCT
 TCTTCCACTTATGTTCCCAAAGGATGAACCTCAACAGAAATCCAACTCTTCTGCTCTGTTAGACAAAGGTACTG
 AGTAGCATCCAGGAAGGGACAGTCCCTGACAAACAAGCTCAGCAAGGGGGAAGGAAGAGCTGAACAAGACCCCTC
 CCAAACTCTCAGGTAGTCAACCAACAGCAAGGCCCCCAACACAGACACATCTGGAAGCTGTGCGCTCCATGGAG
 GCGACTGGGGGCGAGAAAGTCTCAGCAGCCCATCAAGGGCCCAAGAGGGGACATCTAGACAGAACTTACACAAG
 CATTTTGACATCAATGAGCATTGCGCTGGATGATTGTGCTTTTCTGCTGCTGGTGTCTTGTGTGATTGTGGTG
 TGCAGTATCCGAAAAAGCTCGAGGACTCTGAAAAAGGGGCCCGGGCAGGATCCAGTGCCATTGTGAAAAAGGCA
 GGGCTGAAGAAATCCATGACTCCAAACCCAGAACCGGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGAT
 ATCTCTGAAGCTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCTACTAGTTTCTTTGCAATGCCAGTGAG
 AGGGAGGTGTGCTGCTTTCTCCAATGGGTACACAGCCGACACAGAGCGGGCTACGCAGCTCTGCAGCAGCTGGACC
 ATCCGGGGCCCCGAGGCCAGCTCGCCAGCTAATTAGCGCCTTGCGCAGCACCGGAGAAACGATGTTGTGGAG
 AAGATTGTGGGCTGATGGAAGACACACCCAGCTGGAAGCTGACAACTAGCTTCCCGATGAGCCCGCCGCG
 CTTAGCCCCGAGCCCATCCCCAGCCCCAACCGGAAACTTGAGAAATTCGCTCTCTGAGCGTGGAGCTTCCCCA
 CAGGACAAAGAACAGGGCTTCTTGTGGATGAGTCGGAGCCCCCTTCTCCGCTGTGACTCATACCTCAGCGCTCC
 TCCGCTGAGCAGGAAACGGTCTTTATTATCAAAAGAAAGAGGACACAGATGTGTGGCGCAGGTAACGCTGGAC
 CCGTGTGACTTGCAGCTATCTTTGATGACATGCTCCACTTTCTAAATCTGAGGAGCTGCGGGTGTGGAAGAG
 ATTTCCCGAGGTGAGGACAAACTAGACCGGCTATTGGAATTTAGAGTCAAGAGCCAGGAAGCCAGCCAGACC
 CTCTGGATCTGTGTTATAGCCATCTTCGTGACCTGCT**AG**AAACATAGGAGATACTGATTTCTGGAATTTACTCA
 ATTTAGTGGCAGGGTGGTGTTTTAAATTTCTCTGTTTCTGATTGTTGTTGGGCTGTGTGTGTGTGTGTTGT
 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTAAACAGAGANTATGGCCAGTGTGAGTCTTTTCTCTCTCTC
 TCTCTCTCTTTTGTGTTTAAATTAACCTCTTCGGGAAGTTGTTTATAAGCTCTTAAAGCTTTCAGAGTGTAGTGA
 ATACCCACCACATAAGTTTTTTAACTTCCATATTTTCTCCATTTTGCCCTCTTAGTATTTCTCAAGATTAATCTG
 TGCACTTTAAATTTACTTAACTTACCATAAATGCAAGTGTGACTTTTCCCAACACATGGATTGTGAGGCTCTTAAC
 TTTCTTAAAGATATAATGSCATCTGTGAATCTATAAGCAGTCTTTATGTCTCTTAACATTCACACCTACTTTTT
 AAAAACAAATATTATTAATAATTTTATTATTGTTGCTTTATAAAATTTTCTTAAAGATTAAAGAAAAATTAAGA
 CCCCATAATGTTACTGTGAATGCAATCAACTTTGAGTTATCTTTAAATATGCTCTGTATAGTTTCATATTCAATG
 CTGAACCTTGACCAACACTATGCTGATTGTATGGTTTTCACTGGACACCGTGTAGAAATGCTGATTATCTGTAC
 TCTCTTATGCTAATAGTCTTGGGCTGGAGAAATGAAATCTCAAGCCATCAGGATTGTCTATTTAAGTGGCTTT
 GACAACTGGGCCCCAAAGAACTTGAACCTTCACTTTTATGAAATTTAGCTGTCTGGAACACATTTGCTGCACTTT
 GGAAGTCAAAATCAAGTCCAGTGGCGCCCTTTCCATAGAGAAATTTGCCAGCTTTGCTTTAAAGATGTCTCTG
 TTTTTTATATACACATAATCAATAGGTCCAATCTGCTCTCAAGGCCTTGTGCTGTTGGGATTCCTTCAACAAAT
 ACTTTAAATTAACAAATGCTGCAACTGTGAAGAACCTGTCTGATATATTGCAACTATGCTTCCAAATTAACAAATG
 TACTTCTTAATGTCTCAGTTGCCAGGTTCCAATGCAAAAGGTGGCGTGGACTCCCTTGTGTGGTGGGGTGTGGG
 GTAGTGGTGAAGGACGATATACAGAAAAAGTCCCTCAAGTGATCAATTTATTAATAAACATTAAGTGTGTTGA
 AAAAAAAA

ATGGGAAGCCAGTAAACACTGTGGCCTACTATCTCTTCGGTGGTGCCATCTACATTTTGGGA
CTCGGGAAATGTAGGTTAGAGTGGAGGCGGAGCGGATGTACAGAGTCTCGTAATATGTCAC
CATGGGGGAAATGATCCGCTGCTGTGTGAAGCCCTTCTCATTCGATCGATCTTTGGCC
TTGATGATTTGAAATAAGTCTGTGTGCACAGATGCAGATGCTGTTGCTGCACAGATCTCTG
TCACTGCTGCCATTTGAAGTTTTCCTCAATCATCGTCAATTTGGGATCATTTGCATTGATTAGC
ACTGGCCATTTGGTCTGGGCATCCATCTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTGATCGAGCTGATAGCTCGATGTGCAGGAGTCTCGGATGTCAAAGACGGGGAGGA
GAGTACCGCTGTGTCCGGTGGGTGGTCAAGATGCGGTGCTCCAGGTGTTACAGCTGGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAATGTTTGCTGTGCCCAAC
TGGGTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTGCTGGAGGGGCGAGTTC
CGGAGGAGTGTGTGTCATCGATCACTCTTGCCAGATGCAAGGTGATGTCATTATCAACA
CTCAGTATATGTGAGGGAGGAGTGTGCCCTGTGGCAGCTGTGTTACCTTGCAGTGCACAGCCT
GTGCTCATAGAAGGGGCTACAGCTACGCATCGTGGGTGGAATACGTCTTGCTCTCGCAG
TGGCCCTGGCAGGCGACGCCTTCAGTTCTCAGGGCTACCACCTGTGCGGGGGCTCTGTTCAC
GCCCTCTGGGATCATCATCTGTGCACATCTGTGTTTATGATCTGTACCTCCCCAGTATGGA
CCATCCAGGTGGGTGTAGTTTCCCTGTGTGGACAATCAGCCCCATCCCATTTGGTGGAGAG
ATTGTCTACCAACAGCAAGTACAAGACCAAGAGGCTGGGCAAGTGCACATCGCCCTTATGAAGCT
AGGCGGGCCACTCAGGTTCAATGAATGATCCAGCCTGTGTGCTGCCCACTCTGAAGAGA
ACTTCCCAGATGGAAAAGTGTGTCTGAGCTCAGATGGGGGGCCACAGAGATGGAGGTGA
GCCCTCCCTGTCTTGAACACCGCGCCGCTCCCTTTGATTTTCAAACAAGATCTGCAACACAG
GGAGCTGTACGGTGGCATCATCTCCCCCTCATGCTCTGCGCGGGTACTCTGACCGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCCCTGTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCCAGCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCGTGGGCTGTACAC
CCGTGTCACTCTCTTCTGGATGTGATCCACAGCAGATGGAGAGAGACTTAAACCT**TGA**A
GAGGAAGGGGACAGATGAGCCACTGAGTTCCTGAGGTGATGAAGACAGCCGATCATCTCCCT
GGACTCCCGTGTAGGAAGCTGCACACGACGACACACCTTGAGCTCTGAGTTCCGGCACCA
GTAGCAGGCCCGAAAGAGGCCACCTTCCATCTGATTCCAGCACACACCTTCAAGCTGCTTTTT
GTTTTTGTGTTTTTTAGGTTGGAGTCTCGTCTGTGTGCCACAGCTGGAGTGACGTGGCGAAA
TCCCTGCTCACTGCAGCTCCGCTTCCCTGGTTCAAGCGATTCTTGTGCTCAGTCTCCCCA
GTAGCTGGGACACAGGTGCGCCGCCACACCAACCAATTAATTTTGATTTTTATGAGAGAC
AGGGTTTACCATTGTGGCCAGGCTGCTCTCAAAACCCTGACCTCAATGATGTGCTGCTCT
CAGCCTCCACAGTGTCTGGGATTACAGGCATGGGCCACCACCGCTAGCCTCAGCTCCTTTC
TGATCTTCACTAAGAACAAAGGAAGCAGCAACTCTGCAAGGGCGAGTTTCCCATCTGGTTCAT
CTGTTTTCTCTCCAGGCTTCTGCAAAATTCCTGACAGATAGCAGTATTGTGACTCATCAG
TGCAAGAGCCACCAACAGCCACTCAGAAAGACGCACACCGCCAGAAGTGTGCAAGATCGAGTC
ACTGCACGTTTTTCATCTCTAGGACACGAACCAACCCCTTCTACTTCCAAGACTTAT
TTTCATGATGGGGAGGTAAATCAGGAATGACTCGTTTAAAGCTATTTTTCATGATTTCTT
TGTAGCAATTTGGTGTGTGACGTATTATTGTCCCTTGATTTCAAATAATATGTTTCTCTCCCT
CATTTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAATCAATCATCCATGAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISFPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKGDEYRCVVRVGGQNAVLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIHLLPDDKVTALHH
SVYVREGCASGHVVTLQCTACGHRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAAHCVDYDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAVPLISNKICNHR
DVGGGIISPSMLCAGYLTGGVDSQCQDGGPLVCQERRLWKLVGATSFYGICAEVKNKPGVYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACGCGTCCGCTCTAGTCCCGGGCCAACTCGGACAGTTTGTCTATTTATTGCAACGGTCAAGGCTGGCTGT
 GCCAGAACGGCGCGCGCGCGGCACGACGCGCACACACGCGGGGAAACTTTTAAAAAATGAAGGCTAGAAGA
 GCTCACGGCGGGCGCGGGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCGAGAAATCCCTCCGGTCGCGA
 CGCCCGGGCCCGGCTCGGCGCCCGCGTGGGATGGTGACGCGCTCGCCCGCGGGCCCGAGAGCTGCTGCACTGAAG
 GC CGGGACAGTGGCAGCGCGCGCGCTGCCGCTGTCCCCCGCCCGGCCCTCTGCTCGCCCTGGCCGGTCTCT
 GCTCGCGCCCTGCGAGGCGCGAGGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTCAGTGCTCTGT
 TCCGATGCGGACCTCTGATCCGATGGAAGAGCTTCGACTCCAAGATCATCAGAAGTGCTGAATATTGCACT
 ACACCGGGAAGCAAGAACTGATCATAAATCTGGAAGAAATGAAGTCTCATTCGACAGCAATTCA CGGAAC
 CCATCATCTGCAAGACGGTACTGATGTCTCTCCCTCGCTCGAAATTACACGGGTCACTTACTACCATGGACATGT
 ACCGGGATATTCTGATTACAGCACTGCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAATAATGA
 AAGCTATGTCTTAGAACCAATGAAAGTGCAACCAACAGATACAACTCTTCCACGCGAAGAGCTGAAAGCGGT
 CCGGGGATCATGTGGATCACATCACACACCAAACTCGCTGCAAGAAATGTGTTTCCACCACTCTCAGAC
 ATGGGCAAGAGGCAATAAAGAGAGACCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCG
 AGAGTTTCAGAGGCAAGGAAAAGATCTGGGAAAAGTTAAGCAGCGAATTAATAGAGATTGCTAATCACGTTGACAA
 GTTTTACAGACCACTGAACATTCGGATCGTGTGTGAGGCGTGGAAAGTGTGGAATGACATGGACAATGCTCTGT
 AAGTCAGGACCCATTACACAGCTTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAATCCCA
 TGCAATGCGCAGCTTGTCAGTGGGGTTTATTTCAGGGGACCCATCGGCGATGGCCCCAATCATGAGCATGTG
 CACGGCAGACCACTTGGGGGAATTGTCAATGGACCATTACAGACAATCCCTTGGTGCAGCGGTGACCTGGGACA
 TGAGCTGGGGCCAAATTTGGGATGAATCATGACACACTGGACGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA
 AGGAGGCTGCATCATGAACGCTTCCACGGGTACCCATTTCCATGTTGTTCAGCAGTTGCGACGGAAGGACTT
 GGAGACCAAGCTTGGAGAAAGGAATGGGGGTGTGCTGTTAACTGCCGGAAGTCAGGGAGTCTTTCGGGGGCCA
 GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG
 CTGCAATGCCAACCTGTACCTTGAAGCCGAGCGTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT
 GAAGCTGCGAGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACTCCAGAGATTCTGCACAGGGGCCAGCCC
 TCACTGCCAGCCAAATGTACTCTGCACGATGGGCACTCATGTCAAGATGTGGACGGCTACTGCTACATGGCAT
 CTGCCAGACTCACGAGCAGCAGTGTGTACGCTCTGGGGACCAAGTGTCAAACCTGCCCTGGGATCTGCTTGA
 GAGAGTCAATTTGCGAGTGATCTTATGGCACTGTGGCAAAGTCTCGAAGAGTTCCTTTGCGAAATGCGAGAT
 GAGAGTCTAAATGTGGAAAAATCCAGTGTCAAGGAGTGCAGCGCGCCAGTCAATTGGTACCAATGCGGTTTC
 CATGAGAACAAACATCCCTCTGCACGAAGGAGGCCGAGATTCTGTGCCGGGGGACCCAGTGATCTTGGCGATGA
 CATCGCGAGCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGCAAGTATGCTGCTGAATCTGCAATGTAA
 AATATTAGTGTCTTTGGGGTTACAGAGTGTGCAATGCAGTGCCAGCGAGAGGGGTGTGCAACACAGGAAGAA
 CTGCCATCGGAGGCCCATCTGGGCACCTCCCTTCTGTGCAAGTTTGGCTTTGGAGGAAGCAACAGACGGGCC
 CATCCGCGAAGCAGAACCAAGCGAGGAGCTGCAGAGTCAACAGGGAGCCCGGGCCAGGGCCAGGAGCCGCTGGG
 ATCGCAGAGAGTGGCTTACTGCTCTACTGCTCTACTGCACTCATCTGAGCCCTCCCATGACATGGAGACCTGACAGT
 CTGCTGCAGAGGAGTGCAGCGTCCCCAGGGCTCTGTGACTGGCAGCATTGACTCTGGCTTGGCATCGTT
 TCCATGACAAACAGACACACACAGTCTCTCGGGCTCAGGAGGGGAAGTCCAGCTACACAGGCAGCTCTGCAGAA
 CAGTGCAGGGAAGGGCAGCGACTTCTTGGTTGAGCTTCTGCTAAACATGGACATCTGCTGCTGCTCCTGAG
 AGAGTAGCAGGTTACCACTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAGTCTTGGCTTTC
 ACTGAGCCTCCACAGCAGTGGGGGAGAAAGCAAGGGTTGGGCCAGTGTCCCTTTCCCCAGTGACACTCTCAGCT
 TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTATGACATTTATATATGAAAAT
 AGCAGGGTTTATGTTTAAATTTTACAGAGACTTCCGCAACCATTCATCTCATCCAGCAAACTGAATGSCAA
 TGAACCAACTGGGAGAAAGGTAGGAGAAAGGGCGGTGAACCTCTGGCTCTTGTGGACATGCGTGACACAGC
 AGTACTGAGGTTTGAAGGTTTCGAGAAAGCCAGGGAACCCACAGACTCACCACCTTCATTAAACAGTAAGAA
 GTTAAAAAGTGAACCAATGTAAGACCTTACTCATTCCCCGTGGCCATTACTGCATAAAAATAGAGTGCAATTT
 GAAAT

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVSASVRSGLWIPVKSFDISK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCYHGHVVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFFAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETCLKATKYVELVIVADNREFQRQGDLEKVKQRLIETIANHVDKF
YRPLNIRIVLVGVEVWMDMKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFNGMNHDTLDRGCSCQMAVEK
GGCIMNASTGYPPFMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKCGNRFVEEGEE
CDCGEPEECMNRCCNATTTCTLKPDCAVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDVGICYNGICQTHEQQCVTLWGPAGKAPAGICFERVNSAGDPYGN
CGKVSXSFFAKCEMRDAKCGKIQCGGASRPVIGTNAVSIETNIPLQQGGRIILCRGTHVYLG
DDMPDPGLVLAGTKCADGKICLNRCQCNISVFGVHECAMQCHGRGVCCNNRKNCHCEAHWAPP
FCDKFGFGGSTDSGPPIRQAEARQEAAESNRERGQGEPEVGSQEHASTASLTLI

Signal peptide:

amino acids 1-28

FIGURE 31

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTTTGCATTGTTTTCCCTGACAACGAAA
ACAAAACAGTTTTTGGGGGTTT CAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGCAGAAACA
GTGCAAGGAAGGGCAGGANTTCCTGGTTGAGNTTTTTGNTAAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTTTGGCCTTTCACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTT
GGGCCCAGTGTCCCCTTTCCCCAGTGACACCTCAGCCTTGGCAGCCCTGATAACTGGTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTTAGCATTTATTATATGAAAATAGCAGGGTTTT
AGTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

1000
900
800
700
600
500
400
300
200
100
0
1000
900
800
700
600
500
400
300
200
100
0

FIGURE 32

CATCCTGCAACATGGTGAACACACGCCTGGCTAATTTTGGTTGATTTTGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCGCCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAAACTTTTAAAGAAGTTAAT
GAAACCATACCTTTTACATTTTAAATGACAGGAAAAATGCTCACAAATAATGTTAACCACAAA
TTCTGGGATACAAAAGTACAATCTTTACTGTGTAATAACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCTGTGCTAACAA
CTTTTAAACAAAAAATTGCGATCACTTTTAAAGATCAAGAAAAAATTCTGAAGGTCATATGGG
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAGATTGGAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCAGGCTTTCTAAGCAACTTAAATGTTTGTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGAGTGATGTCACTGCCATTATGATGCCCC
TTGAATATAAGACCCTACTTGCTATCTCCCTGCACCAGCCAGGAGCCCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGGATGGTGGCGA
TGCTCATTCTGGGCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTGT
CCTTTAACAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAAACAGTGCCT
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCATACACTAAAACCGAATCCAAATGGTGCCT
AGAAGTTCATGTGGCAAGGAAAAAACAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCTGTACTCCCAGTACTTTGGGAGGCCATAGCCGGCGGATCACCTG
AGGTGAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAAGCCCATCTGTAATAAAAAACA
AATATTGACTGGGCGTGGTGGTGAGTGCTGTGATCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACCTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCTGTAATCCCGGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTGAGGATCA
AGACCATCTGGCTAATACAGTGAAACCTGTCTCTACTAAAAATACAAAAATTAGCCGGG
GATGGTGCAGGCACCTGGAGTCCAGTACTCGGAGGCTGAGGCAGGAGAAATAGCGTGAA
CTCAGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309
```

```
><subunit 1 of 1, 67 aa, 1 stop
```

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALOPTAGLLVVLLAL

LHLYH

Signal peptide:

amino acids 15-27

Figure 1 consists of 12 line graphs, labeled (a) through (l), arranged in a 6x2 grid. Each graph plots the bacterial concentration in log₁₀ CFU/g on the y-axis (ranging from 0 to 10) against time in hours on the x-axis (ranging from 0 to 24). The graphs show the effect of different heating treatments on the growth of *E. coli* O157:H7 in ground beef. Graph (a) is the control, showing a steady increase in bacterial load from approximately 10^{1.5} to 10^{7.5} CFU/g. Graphs (b) through (l) show the effect of heating at 100°C for increasing durations: 15 min, 30 min, 45 min, 60 min, 90 min, 120 min, 150 min, 180 min, 210 min, 240 min, and 270 min, respectively. As the heating duration increases, the initial bacterial load decreases, and the rate of growth slows down, with some samples reaching a lower steady state or showing a slight decrease over time.

FIGURE 34

GCGCGCGCGAGAGCGCGCCGACGCCCGCGCGATGCGCGCGCGCCGAGGACGCTCTCCGCTGCTGCGCCGCG
 CGGCGGCGCTGACTGCGCTGCTGCTGCTGCTGCTGCGGCATGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCCGAG
 AGGCGCGCGCGCGCGCGCGCGGACGCGGCCCGCGCGCGAGACGCGCGAGGACGGACAGGACCCGCAACAGCAAC
 TGTCACAGCGCGCATGTTACGCAACGCGGATCCAGAGCGCGCGCACCTTCGTCACTGTTCTTCGCGCGCTTGCTGCT
 GACATGCGCAGCGCTCGACCGCATCTGGAATGACCTGGGAGACAATACACACAGCATGGAAGATGCCAAAGTCT
 ATGTGGCTAAAGTGGACTGACGCGCCCACTCCGAGCTGTCTCCGCCAGGGGGTGGAGGATACCCACCTTAA
 AGCTTTTCAAGCGAGGCGAAGAGCTGTGAAGTACACAGGGTCTCTCGGAGCTTCCAGACACTGGAAGAACTGGATGC
 TGCAGACACTGAAACGAGGAGCCAGCTGCACACAGAGCGGAAAGTGGAAACGCGCGAGTGGCCCCGAGCTCAAGCAAG
 GCGTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCACGTTGCACAAAGCGACCACTTTATCAAGTTCTTTCGCTC
 GACTGTGTGGTCACTGCAAGCGCTGGCTCCAACTCTGGGAGCAGCTGGCTCTCGGGCTTGGAGATTCGGAATCTCGAAACTG
 TCAAGATTGGCAAGGTTGATTGTACAGCACTATGAACCTCTGCTCCGGAAACAGGTTCTGTGCTATCCCACTC
 TTCTCTGGTTCCGAGATGGGAAAAAGGTGGATCAGTACAAGGGAAAGCGGGATTGGAGTCACTGAGGGAGTACG
 TGGAGTCGACGCTGCAGCGCACAGAGACTGGAGCGACGAGAGCCGTCACGCGCTCAGAGGCCCCGGTGTGGCAG
 CTGAGCCCGAGGCTGACAAGGGCACTGTGTGGCACTCACTGAAAAATAACTTCGATGACACCATTGCAAGAGAA
 TAACTTTCATCAAGTTTATGCTCCATGGTGTGGTCAATTGTAAGACTCTGGCTCCTACTTGGGAGGAACCTCTCTA
 AAAAGGAATCCCTGGTCTGGCGGGGGTCAAGATCGCCGAAGTAGACTGCACTGCTGAAACGGAATACTGCGAGCA
 AGTATTCGTAACGAGGCTACCCCAAGTATTGCTTTCCGAGGAGGGAAGAAAGTCAGTGAGCACTGGAGGCA
 GAGACCTTGACTCGTTTACACCGCTTTGTCTGAGCCAAAGCGAAAGACGAACCTTAGGAAACACAGTTGGAGGTCA
 CTCTCTCTGCGCAGCTCCCGCACCTTGCGTTTAGGAGTTCACTCCACAGAGGCCACTGGGTTCCAGTGGTGGCT
 GTTCAGAAAGCAGAACTACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTCCAAAGCCAAACACTCTACAG
 ATTCTTTATTAAGTTAAGTTTCTCTAAGTAAATGTGTAACCTCATGCTCACTGTGTAAACATTTTCAGTGGCGATA
 TATCCCTTTGACCTTCTCTGTATGAAATTTACATGGTTTCCCTTTGAGACTAAAATAGCGTTGAGGGAATGAAA
 TTGCTGGACTATTTGTGGCTCTGAGTTGAGTGATTTTGTGTAAGAAAGCACATCCAAAGCATAGTTTACCTGC
 CCACGAGTTCTGGAAGGTTGGCTTGTGGCAGTATTGACGTTCTCTGATCTTAAGGTACAGTTGACTCAATC
 TGTGTGTGCTGCTAGCATGGAGCAGATTGAAATGCRAAAACCCACACTCTGGAAGATACCTTCACGCGCGCTGC
 TGGAGCTCTGTGCTGTGAATACCTTCTCTCAGTGTGAGAGGTTAGCGGTGATGAAAGCAGGTTACTTCTGACC
 GTGCTGAGTAAAGAAATGCTGATGCCATAACTTTATGTGTGATACCTTGTCAAATCAGTTACTGTTCAGGGAT
 CCTCTCTGTTTCTGCGGGGTGAAACATGCTTTTAGTTCTCTCATGTGTAACAGAAAGCAGGTTACTTCTGACC
 TGGATGTCTTCTTAGAAAGGCTAGGCAATGGAATTTCCACGAGGCTCATTTCTCAGTATCTCATTAACCTATTGA
 AAGATTCAGTTGATTTGTCACTGGGGTGAACAGACAGAGGCTTTCCAGGCGCTGGGATTCACAGGAGG
 TCTCGAGCCTGCTGAAGGCGCTTAACTAGAGTTCTAGAGTTCTGATTCGTGTTCTCAGTAGTCTTTTAGAGG
 CTGCTATACCTGCTCTGCTTCAAGGAGGTCGACCTTCTAATGATGAAGAAATGGGATGATTTGATCTCAAGAC
 CAAGACAGATGTGAGTGGGCTGCTCTGGCCCTGCTGTGCAACGGCTGTGGCAGCTGTGATGCCAGTCTCTCTA
 ACTCATGCTGTCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATCAGGCTTAAAGCTCTAAGATA
 GATAGGTGTTTGTCTTTTACCATTGAGCTACTTCCCATATAACCACTTTGACATCAACACTCTTCAACCACT
 CCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACCTGGTGTAGGAATCTTAGAAACAGAGCACTTATA
 CTGTCTGTCTGAGGCGAAGATAACAGCAGCATCTCGACCAAGCTCTGCTCTTAAAGGAAATCTTTATTAATCAG
 TATGGTTACAGATAATTTCTTTTTTAAAAAAACCAACCTCTCAGAGAGCACAACCTGTCAAGAGCTTTGTACA
 CACAACCTCAGCTTTGATCAGAGTCTGTATTTCAAGAAATCAAGTGGTACAAATTTGTTGTTTACACTAT
 GATACCTTCTAAATAAATCTTTTTTTTAA

10016177.102501

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARFGRLLPLLARPAALTALLLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS
KHLTYADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNLDGKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRGYPTLKLFKPGQEAVKYQGPRDFQTLLENWMLQTLNEEPVTPEPEVEPPSAPE
LKQGLYELASANFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHVELCSGNQVRGYPTLLWFRDGGKKVDQYKGKRDLESLEEVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLFRRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

Signal sequence:

amino acids 1-32

FIGURE 36

CTTTTCTGAGGAACCACAGCAATGAATGGCTTTGCATCCTTGCTTGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACAATTTACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGAATCCCTAACCCACTGCAGGATTTCGG
GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTTCGGGTGTTTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT
ACGTATTTGCTATTTTCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTCTCTG
ATTGTACTACATTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLQLIQLGLDIDSRPTAEVCATHTTISPGPKGDDGEKGDPEEG
KHGKVGRMGPKGIKGEIGMDQGNIGKTGPIGKKGDKEKGLLGIPGEKKGAGTVCDGRY
RKFGVQLDISIARLKTSMKFVKVNIAGIRETEEFYIYVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDECHLTMYFVCFEIKKKK

Signal peptide:

amino acids 1-25

FIGURE 38

GGTCTCATCGATTGGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCGTGTGGAGTGGTGGGGGCTGGGTGGGAATCGGCGTGT
GCCAGCGCACGCGCTCCTCGGAAGGAGAAGTCTCAGCTAGAAACGAGCGGCCCTAGGTTT
CGGAAGGGAGGATCAGGGATGTTTGCAGCGGCTGGAACACAGACGGTGGCGATAGAGAAAGC
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCCCTGCTGCTGTTGCTACCCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCGAGTTGCGCTGGCTTCGGCCGGAAGTGGCCCTTGGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCCGAGGGGGGCTGCAGCCTGGCCCTGGCCGCTCGCGGAAGTGGCCACGAGCGCGCCGCGC
ACACCTTTCTCATTCAACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCCGAGAGTAAC
AGGGCTGCACGCGCTTCTCCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGCGCGGACAG
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATCAGCGGCCG
GAAGCGGCGCGGAGTTTTCGCGGAGGGGACGGTGCCGCCAGAGGTGGAGGAGCCGCCGCCCT
CTGTCACTGGAGCACTGTGGCGCTGCTCCTCCCCGCTGGCCAGAGTTTCTGTGGCTCTG
GTTGGGCTGGCCAAAGGCGGGCTGCGCACTGCCTTTGTGCCACGCGCTGCGCCGGGGCC
CCCTGTGCACTGCCTCCGAGCTCGCGCGCGCGCGCTGGTGTGCGCGCAGAGTTTCTG
GAGTCCCTGGAGCCGACCTGCCGCCCTTGAGAGCCATGGGGCTCCACCTGTGGGTGCGAGG
CCAGGAACCCACCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTCTTCCCCCAGAGCATAAACAGACAGCTGCCTGTACATCTTC
ACCTCTGGCACCACGGGCTCCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCAC
TCTACCACATGTCCGGTTCCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCAAGTTCTGGGAAGATTGCCAGGACGACAGGGTGAC
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCGGAGCAAGGCAG
AACGTGGCCATAAGGTCGGCTGGCAGTGGGCAGCGGGCTGCGCCAGATACCTGGGAGCGT
TTTGTGGCGCGCTTCGGGGCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCTGGCTTTACAAGC
ATATCTTCCCCCTCTCCTTGATTGCTATGATGTCAACACAGGAGAGCCAATTTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCGGTAAGCCA
GCAGTCCCCATTCTGGGCTATGCTGGCGGGCCAGAGCTGGCCACAGGGGAAGTTGCTAAAGG
ATGCTCTTCGGCTCGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAGAATGTGGC
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACCTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCTGCCCTCCC
CAGCTTTGGACCTTATGCAGCTCTACACCCACGCTGTCTGAGAACTTGCCACCTTATGCCCG
GCCCGGATTCCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAACAGCAGAGAAG
TTCGGATGGCAAAATGAGGGCTTCGACCCAGCACCCTGTCTGAGCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCTTACCTGCCCTCACAATGCCCCGTACAGCGCCCTCCTGGCAGGAAA
CCTTCGAATCTTGGAAACTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGGG
CCGTTGTCAGGTGTACTGGGCTGTGAGGGATCTTTCTATACCAGAAGTCCGGTCACTATTTT
GTAAATAATGTGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAATAAAAAAAAAAAAA
AAAAAATAAGGGCGGCGCGACTCTAGAGTCAGCTCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCATGGCCCACTGTTTATTATGCAG

10016177.102501

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913

><subunit 1 of 1, 730 aa, 1 stop

><MW: 78644, pI: 7.65, NX(S/T): 2

MGVCQRTTRAPWKEKSQLERAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLLPL
LLLLKHLWPQLRWLPADLAFVRLCCKRALRARALAAAAADPEGPEGGCSLAWRLAEQAQQ
RAAHTFLIHGSRFFSYSEAESNRARAFLRALGWDWGPDPGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAPLSPGATVALLLPAGPEFLWLWFGGLAKGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTTGEPIRDPPQGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPQLAQK
LLKDVFRPGDVFNTGDLVLCDDQGFLRFHRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYPARPRFLRLQESLATTTETFK
QQKVRMANEGFDPSTLSDPLYVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
 CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTC
 GCTTGTCCATCTCCCTCCCGGGGAGCGCGCGCGCTCCCACCTTTGCCGCACACTCCCGGC
 GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGAATTCGACGCTCTGAACCC
 CCATGGTGGTTTTTTAAACACTTCTTTTCTCTCTTCTCGTTTGGATTGCAACCGTTTCCA
 TCTGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCAGCCAGCCCTTGTGGCTTGCCATCGT
 CCATCTGGCTTATAAAAAGTTTGTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCTCGG
 CTGGCAGAGGGGGTGACGCTGGGCAGCGCGCAGGAGCGCGCCCTGCTCTGGCGCGCTTT
 CGGTTTGGGGCGAAGGTGAAGAGCGCACCGCGCTGGGGTTTACCGAGCTGGATTGTATG
 TTGCACCAATGCTTCTTGGATCGGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
 TCCCCGCGGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCGCCAGGCGTACGGTGCC
 AAGGATTACGCTTGGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
 TCCTCAGGAATATACATGCTGCACCAACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAC
 TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCCTTTTGTGTCACGG
 CATAAGAAATTTGACGAATTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
 TATGTTTGCACGACCTATGGCATGCTGTACATGCAGAATTCAGAGATCTTCCAGGACCTCT
 TCACAGAGCTGAAAAGGTACTACACTGGGGTAAATGTGAATCTGGAGGAATGCTCAATGAC
 TTTTGGGCTCGGCTCCTGGAACGGATGTTTACAGCTGATAAACCTCAGTATCACTTCAGTGA
 AGACTACCTGGAATGTGTGAGCAATACACTGACCAGCTCAAGCCATTGGAGACGTGCCCC
 GGAAACTGAAGATTGAGTTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG
 ACTGTGGCGCAGAGAAGTTGCAAAACGAGTTTCCAAGTCAAGCCAAACCCAGGGTGTATCCG
 TGCCCTCATGAAGATGCTGTACTGCCCATCTGTGCGGGGCTTCCACTGTGAGGCCCTGCA
 ACAACTACTGTCTCAACGCTCATGAAGGGCTGCTTGGCAAAATCAGGCTGACCTGCACACAGAG
 TGAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAAAT
 TGAGTCGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
 ACAGCATGCAGGTGTCTGCAAAAGGTCTTTTCAAGGATGTGGTCAGGCCAAACCTGCTCCAGCC
 CTCAGATCTGCCCGCTCAGCTCCTGAAAATTTTAAATACAGTTTTCAGGCCCTACAATCCTGA
 GGAAAGACCAACAACTGCTGCAGGCACAAGCTTGGACCGGCTGGTCACAGACATAAAAGAGA
 AATTGAAGCTCTCTAAAAAGGTCTGGTCAGCATTACCTTACACTATCTGCAAGGACGAGAGC
 GTGACAGCGGGCAGCTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
 GCCTGAGATCATGAATGATGGGCTCACCAACAGATCAACAAATCCGAGGTGGATGTGGACA
 TCACCTCGGCCTGACACTTTTATCAGACAGCAGATATGGCTCTCCGTGTGATGACCAACAAA
 CTAAAAACGCTCATCAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
 CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCACGGAGTTTGTAGTTTGTACCA
 CAGAGGCCCGCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCAGCGTGGC
 CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCTGGCACTGCAGAGACTGTGCGAGATA
 ATCTTTGGGTTTTTGGTCAGATGAACTGCATTTTAGCTATCTGAATGGCCAACCTCACTTCTT
 TTCTTTACACTCTTGGACAATGGACCATGCCACAAAACTTACCCTTTTCTATGAGAAGAGAG
 CAGTAATGCAATCTGCCCTCCCTTTTGTTTTCCCAAAGAGTACCGGGTCCAGAGCTGAACCTG
 CTTCTCTTTCTTCTTCACTATCTGTGGGACCTTGTTTTATTCTAGAGAGAATTCTTACTCAA
 ATTTTCTGACAGGAGATTTCTTACCTTCACTTGTCTTATGCTGCACAGTAAGGAAT
 CTCACGTTGTGAGGGTTTTTTTTTCTCATTTAAAT

10016177-102514

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADVKA~~R~~SCGEV~~R~~QAYGAKGFS~~L~~ADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTTFVSRHKKFDEFFRELLENAEKSLNDMF
VRTYGM~~L~~YMQNSEVFQDLFTELKRYYTGGNVNLEEMLNDFWARLLERM~~F~~QLINPQYHFSE~~D~~Y
LECVSKYTDQLKPFQDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNQENSMQVSAKV~~F~~QCGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGH~~S~~KARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDDVCPTEFEFVTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

FIGURE 42A

CGGACCGGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG
CAGTTTGCACGCGCTGCGCCGGTGGCCAACTACGCAAAGACCAAGCGGGCTCGCGCGGACCGGCCGCGGGG
TAGGGACCCGCGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGAGAGT
GGAGGGAGATCAGGAAACGGCTTCTCTCCTCCTCGCCGCTGGTGTGAGTGTGCGGGAGATTGGCAAACGCCCTAGG
AAAGGACTGGGAAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG
TCTGATCTGAGCCGACGCGACGCGTCCACTTCGCACTTCTTCCAGTGTGGGGACCGCAGGACAGACGCGCCGA
TCCCGCCGCCCTCCGTACCAGCACTCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCGCTCTGGCCACGA
AAAGTTCTCTGTCCAATGTGATTCTCAATCTCTGTCTGGTTTTTTTCTCCAGAGAATTTTGGGTGGAGATATTA
ACTTTTTCTTTTTTTTTTCTTGTGGAGAGCTGCTCTAGGGAGGGGGAGGAGGAGGAGAAAGTGAATGTGTC
TGGAGAAGAGCGAGCCCTCCTGTGTTCTTCCGAGTCCCATCATTAAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGTTGACAGCTGAGAGGAGAGGAGGATTCTTGGCAGGTGGAGAGTCTTACCCTGTGTTGGGTGTCATG
TGTGCGCCCCGAGCGCGCGGGCGCGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGAATGCTGCCA
GGGCTGTGCGGGGCATCGCCTCCGCTTCTCCACAGGCCTGTGTCTGCTCTGGAAGAATGCTAGCAATGGGG
CGCTGGCAGGATTCTGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCTGGGGCCAGGCCCTTAGAAGAGGAG
AAGAAGGGGCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCAGCAAACTTCCACCTCCAGGCCCATCTCA
TTTTCTCCTAGCGGATGATCAGGGATTAGAGATGTGGTTACCACGGATCTGAGATTAAAAACCTACTCTTG
ACAAGCTCGCTGCCGAGGAGTTAAACTGGAGAACTACTATGTCCAGCCTATTGTGCAACCATCCAGGAGTCAGT
TTATTACTGGAAGATATCAGATACACACCGGACTTCAACATTCTATCATAGACCTACCCAACCCAACTGTTTAC
CTCTGGACAATGCCACCTCACTCAGAACTGAAGAGGTGGATATTCAACGCATATGTCGGAATAAGTGGCACT
TGGGTTTTTAACAGAAAAAATGTCATGCCACAGGAAGAGGATTGTATCACTTTTTTGTCTCCCTTTGGGAAGTG
GGGATTACTATACACTACAAATGTGACAGTCTTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAACTCTAGCTTCCCAT
ACCCCAAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTCATTACCACCTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACCAGGAGAAGATATGCTCCATGCTTCTCTGCTAGATGAAGCAA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTCTATAACAACAGCATATTATCTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGAGTAATGGCCTCTCAGAGGTAGCAAGGAACATATTGGGAAGGAGGATCC
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTGAAGGAACCTTGTGCACATCACTG
ACTGTACCCCTCACTTCACTTATGCTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTTATGATATCT
GGGAGACCATAAGTGAGGGTCTTCGCTCACCCCGAGTAGATATTTTGCAATAACATTGACCCCTATACACCAAGGC
AAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTCAGCACTG
GAAATTGCTTACAGGAAATCTTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTTACGCAACCTGGGACGGAACCG
GTGGCAATGAACGGATCACTTGTCACTGGCAAAAGTGTATGGCTTTTCAACATCAGCCGACCCCATATGA
GAGGGTGACATCTTAAACAGGTATCCAGGAATCTGAAGAAGCTCCTACGAGGCTCTCACAGTTCAACAAAC
TGCAAGTGGCGTCAGGTATCCCCCAAGAACCCAGAAAGTAACCTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAGCCAGCAAAAATCAGGCTGAGAAAAAGCAAAAGGAAAGCAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGSTAAACAGCAAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
GGCTGTTTTTATGCTGTGCGCATCCAGAGACTTCTGCCCTGGCCGCCACACTGCAAACTGTCTGCTCAGTG
CCAAGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGGAAACGTG
GTGAGTCTGTAGTTCACCTGCTGTGCTTCACTCAACTGACCAAACTGCTTTGAATTATAGGAGGAGAACATA
ACCTACCATCTCGCAAGCATGCTAATTTGATGGAAGTACAGGGTAGCATGATTAAGAACTACCTTTGATAAATTAC

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTGYLSWGQALEEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYPHGSEIKTPTLDKLAATEGVKLENYVQPICTP
SRSQFITGKYQIHTGLQHSIIIRPTQPNCLPLDNATLPQKLKEVGYSTHVMGKWHLGPNRKEC
MPTRRGFDTPFGSLLGSGDYIYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLEAINNVTLA
LKTYGFYNNSIIIIYSSDNGGQPTAGGSNWPLRGSKGTWEGGIRAVGVFVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSRPVDILHNIDPYTPRQKMAPG
QQAMSGTLQSSQPSECTGNCLQEILATATGSPLSLSATWDRGTGTMNGSPCQLAKVYGFS
TSQPTHMRGWTYLTGIQES

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCGCTCTGCCCTGGAGCCTTGCCTGCCGCTGCTGCTCTCTCTGGGTGGCAG
GTGGTTTTCGGGAACGCGGCCAGTGCAAGGCATCAGGGTTGTTAGCATCGGCACGCTCAGCCT
GGGTCTGTCTACTATGGAACATAACTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACTGGATGTAAGTTTGGTGAAGTGCCTGGGACCAACAAT
GCAGATGCTTTCCAGGATACACCGGGAACCTGCAAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACAGATGTGTGAATACACAGGAAGCTACAAGTGCTTTTGCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACCTTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCTCTGTGTCCATCTCAGGA
CTCCGCTGGCCCCAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTGGGAAGCTACTACTGCAATGTCCACA
TTGGTTTCGAACATGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCAAGGGTCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGCAATGGACTTCGGTGTCTGCTATCCCTGAAAATTTCTG
TGAAGGAAGTCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCCAAAAATTAAAAATGTTACCCAGAACCCACAGGACTCC
TACCCCTAAGGTGAACCTTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAAC
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAATGAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGCATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTTCTGGTCCAAAGGAAAGCGCTAATCTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAACAGGATAGAGAAGATGATTTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTCAACAAGAAAGACATTGGCCGATTTGAACCTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGTCTTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTCAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGAAAAATTCAGTTGTATCAAGGAATCATGTC
TACCAAAAGCATCATTTTTGAAGCAGAACGCTGGCAAGGGCAAAACCGGCGAATCGCAGTGG
ATGGCGTCTTGCTTTTTCAGGCTTATGTCCAGATAGCCCTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTGACTTTGTATGTGCTAGTTCCCTGGTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTAGAATTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATATTG
TAAGATGCCCTTTCTGTATAAGATATGCCAATATTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTTCCNCATTATATATAAAATNTGGAAANGTCACTTTATCTC
CCCTCCTCNGTATATCTGATTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTCTTA
GAAAATAGAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACCTCTTGGGA
AACTATGACATCAAGAGTAGACTTTTGCCTAAGTGCTTAGCTGGGCTTTTCATAGCCAAAC
TTGTATATTTAATCTTTGTAATAATAA

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNRNSKGVCE
ATCEPGCKFGECVGPNNKRCRCPGYTGKTCSDVNECGMKPRPCQHRVCVNTGSGYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCDETEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRVCNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCASHANCFTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKLLAHKNSMKKKAKIKNVTPEPTRTPPK
VNLQPFNYEEIVSRGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGCTGTGCGCGCTGCTCCTGCTCTTGTTGTCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCC
AGAAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGCTGTGAGCCAGAAGAGTGCAT
GAGCTGGAAAGGGTGAAAAGAAGATGCCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT
TGTTTTGCCCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCAGCGTTCTCTGTGC
ATGGATACCAGCTTGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTCTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTAAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATGTGGAGAATTCCTTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACCAACCTTTCTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAATC
TTTAAGACAAAACATGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTAAATAGATATGACTTGTCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343

><subunit 1 of 1, 289 aa, 1 stop

><MW: 32268, pI: 9.21, NX(S/T): 0

MVVVVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPTWAWWITNKMGGKRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTGCGCGCTCTTCTCCTCCAGTGGGGCAGGGGTTTC
 GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTCCGG
 TTGCTACTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
 TGATGCTGCTGAGCTCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTC
 TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
 GTGGCTCAGTTTTCCGGAAGGTCCAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
 ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCTGTG
 GACAGCGTGGCCCCGCGCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCCC
 TGGCAGGCGCAGTGTGAGGAGGCAAGGAGCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
 CTGGGTCTCTCACTGCTGCCCACTGCTTTGAAAAGGCAGCAGCAACAGAAGTGAATTCCTGGT
 CAGTGGTCTCTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
 GCTGCCCTGCAGTTGGCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
 GCAGCTGCCCCACCCACGACCCACACACCCCTCTGCTGCCCCAGCCCGCCCATCGCTTCC
 CCTTTGGAGCCTCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGCTCCTGGGACC
 CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAACTGTATCTACAACCAAGCT
 GCACCAGCGACACCTGTCCAACCCGCGCCCGGCTGGGATGCTATGTGGGGGCCCCAGCCTG
 GGGTGCAGGGCCCTGTCAAGGAGATTCCGGGGGCCCTGTGCTGAGCCTCGAGCCTGACGGA
 CACTGGGTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT
 GCTGTGACCAACACAGCTGCTCAGGTTCTCTGGCTGCAGGCTCGAGTTCAAGGGGCGAGCTT
 TCCTGGCCCCAGAGCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
 TCCTTGAGGACAGCAGGTCCCAGGCAGGAGCACCCCTCCCATGGCCCTGGGAGGCCAGGCT
 GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGGCGGTGCTAACTG
 CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCGA
 CCGGAGGAGTGGGGCCTGAAGCAGCTCATCTGTCATGGAGCCTACACCACCTGAGGGGGG
 CTACGACATGGCCCTCTGTGCTGCTGCCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
 TCTGCTGCCCTATCTGACCACCACCTGCTGATGGGGAGCGTGGCTGGGTTCTGGGACGG
 GCCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCCTCTGGGGCCTAG
 GGCCTGCAGCCGCTGTCATGCACTCCTGGGGGTGATGGCAGCCCTATTCTGCCCGGGATGG
 TGTGTACCACTGCTGTGGGTGAGCTGCCAGCTGTGAGGGCCTGTCTGGGGCACCACTGCTG
 CATGAGGTGAGGGGCACATGGTTCTTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
 CCCCGCCAGGCCGGCGGTCTTCAACGCGCTCCTGCCTATGAGGACTGGGTGACAGTTTGG
 ACTGCGAGGTCTACTTCCGCCAGGAACAGAGCCCGAGGCTGAGCCCTGGAAGCTGCCTGGCC
 AACATAAGCCAACCAACCAAGCTGCTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
 AGGCAGGCCAAATGGCATTACTGCCCTGTCTCTCCCAACCTGTCTATGTGTGATTCCAGGCAC
 CAGGGCAGCCCCAGAAGCCCAAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
 CTCCCCACCTGTGAGGACAGGGGTGTCTGTGGACACTCCCCACCCCACTCTGCTACCAAGC
 AGGCGTCTCAGCTTCTCTCTCTTACTCTTTTCAGATACAATCACGCCAGCCACGTTGTTTT
 TGAATAATTTCTTTTTTGGGGGGCAGCAGTTTTCTTTTTTAAACTTAAATAAATTTGTTAC
 AAAATAAAA

100677-16594

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWFWQASVRRQGAHICSGSLVADTWVLTAAHCFEAAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDDLALLQLAHP TTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRLNRLRLIISRPTCNCIYNQLHQRHLSNPAPFGMLCGGPQPGVQGPCGDSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLDPDGERG
WVLGRARPGAGISSLQTVPVPTLLGPRACSRSLHAAPGGDGSPI LPGMVCTSAVGELPSCGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCGGCCCCCATTCGGGCCGGGCCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCAGAGTCCGGCGCGGCGCGGAGGGGAGCCTTCCACCACGGGGAG
CCCAGCTGTGAGCCGCTCACAGGAAGATGCTCGCTCGGCGGGGAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA
GCTGTGTCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC
TCTTCCCGACCTGTGTCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGATTTCGGCAGCGCTGCCGTGAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCAGCATGACCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGTACCAGGGCTACCTGAGGCTGAGGTGTTT
TGGCAGGATGGGAGGGTGTGCCCTGACTGGCAACGTGACCACGTGCGAGATGGCCAACGA
GCAGGGCTTGTGTTGATGTGCACAGCGTCTGCGGGTGGTGTGGGTGCGAATGGCACTACA
GCTGCCTGGTGCACAACCCGTGCTGCAGCAGGATGCGCACRGTCTGTCAACCATCACAGGG
CAGCCTATGACATTCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTTCTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCCTGCGCTGCAATGGGGCTGCAGTGTGAGCCCC
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCAAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTCCAAAGTCATCCTGCTGCCTTTT
TTCTTATAGACACAATGAACAGACACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACACCCTC
TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCAAAGATGCAATATTGAGA
CTGACTGACCCCTGCCTTATTTCAACAAAGACACGATGCATAGTCACCCCGGCTTGTTC
TCCAATGGCCGTGATACACTAGTGATCATGTTTCAGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGC GGCTCAACATCTCTGGAGTCTAGAAGCTGTTTCTTTT
CCCTCCTTCTCTCCCTGCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCACCACCATGGTGTCTATTCTGGGGCTGGGGCAGTCTTTCTCTGGC
TTGCCCTTGCCAGCTCTCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCG
GATGTCTCTCTCTCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTGTAAACTGGGGGTATATTTGGGGAAAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQ
LNLIIWQLTDTKQLVHVSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQRVRVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDQGQGVPL
TGNVTTSQMANEQGLFDVHVS LRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLLVALAFVCWRKIKQSCBEEENAGAEDQDGEGEKSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 53

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194
><subunit 1 of 1, 211 aa, 1 stop
><MW: 24172, pI: 5.99, NX(S/T): 1
MRLFLWNAVLTLFVTSLIGALIPPEVKIEVLQKPFICHRKTKGGDMLVHYEGYLEKDGSL
FHSTHKHNNQPIWFTLIGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKKGAVVNESHHDALVED
IFDKEDEDKDGFIAREFTYKHDEL

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTTCTCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTTAGCAT
CCAAACCATTCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCTCTCCGTAGCCACCCGA
CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCAACAGTAC
CTGCCACCTCAACGTCCTCAATGGCTCTGACGCCCGCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCAACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCACAACTGCCTC
TGAGGAGATGTTCTCAGTTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCACGACAAGTACGATGTGTCTGGTGATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTTACAACTGCTACATCATGAACCCCCCTGACCGCCACCGTGG
CCATGGCAAGATCCATCTGCAGGTCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCTCCGTGGGGGGCTTCTGGCTGTGGTTCATCTTGGTGCTGATGGTG
GTC AAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACCGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCTCTCCGCCCTGTACAGTGACCTGCCTGCTCG
CTCTTGGTGTGCTTCCGTGACCTAGGACCCAGGGCCACCTGGGGCTCCTGAACCCCG
ACTTCGTATCTCCACCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCGCTGGGGAGAGGAGAAAGGGCTCCACCTGCCAGTCCCTGG
GGGAGGCAGGAGGCACATGTGAGGGTCCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA
GGAGGGGCGCTGTACCTGCCAGTGCTTGCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGCTTCTCTGTGCTGACAGCGCTCCCTCAGGAGGGCTTGGCTGGCACGGCTG
TGCTCCTCCCTGCTCCAGCCAGAGCAGCCATCAGGCTGGAGTGAGTGAGTTCCTGA
AACTTGGAGGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGGAAAGCCAGGG
CTGCAGGCAAACTGGACATGTGCCCTGGCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
GCTAGTGGCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGGCTGAGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCTTCTCCATGGGGTAACACCCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTCTTTGGGGCTTCAGTGAGAACTCTCCGAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCG
GCATAAGGGGAGGCCTTGGAACTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG
CTACTCGCTCCTCTCCCAACAACTCCCTTCTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCCCTGTAATCCACGACTTTGGGAGGCCAAGCGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGCAAAACCCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG
GTGGCGTGTGCTGTAAATCCAGCTATTGGGAGGCTGAGGCAGGAGAAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAAGTACGATAGTGCCACTGCAATTCAGCCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKI INLKLERFQDRVEFSGNPSKYDVSVMRLRNVQPEDEGI
YNCYIMNPPDRHRGHGKIHLQVLMEEPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRKK
KEQKLSTDDLKTEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTAGCATCCAACCATCCTCCCTTGTA
GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCACCCNACTAACATCTCAGTCTCTGAA
AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCACCCCTCAACGTCC
TCAATGGCTCTGACGCCCGCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
AGTTCTCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCTCCAG
TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGGTGGAGTTCTCAGG
GAACCCAGCAAGTACGATGTGTGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA
TTTACAAC TGCTACATCATGAACCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCTCAACGTCTCTCAATGGCTTTGACGCCCGCCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACCTGGC
TCTGAGGAGATGTTCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGAAGTTTCTCAGGGAACCCAGCAAGTACGATGTGTCGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTACAAC TGCTACATCATGAACCCCC

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 59

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189
><subunit 1 of 1, 412 aa, 1 stop
><MW: 46658, pI: 6.65, NX(S/T): 4
MGLHLRPYRVGLLPDGLLFLLLLLLMLLADPALPAGRHPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVVPFGKTFSL
EFLDSPKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNRI
PVIGPLKIREQQRSVSTSWLLPYNTWSPEKVVFQTPPTINYTLRDYRKFFQDIGFEDGWL
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQ
AQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLG
```

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGACGCGGCGGACGGCGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCTTGACGACGCGCA
GGTGGTGGCGCGCGCGTGTGCTTGGTCTTCGCCTTGATCGTGTCTCTCGATCTATGGTG
AGGGGTACAGCAATGCCACGAGTCTAAGCAGATGTACTGCGTGTTC AACCGCAACGAGGAT
GCCTGCCGTATGGCAGTGCCATCGGGGTGCTGGCCTTCTGCGCTCGGCCTTCTTCTTGGT
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCTGTGGTTTGTGGTTTCTGCTTCCTCACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCGGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACT
GCCTACGCTCCTACCCAGGTGCATCTGTGGACAACTACCAACAGCCACCCTTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTAGTGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCTCTGCCCTGGACTTTCCCATCAGCCTCTCGAACTGCCA
GCCCTCTCTTTCACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCTGTGCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTT TAGCTAGTGTTTTCTCGCTTTTAATGA
CCTCAGCCCCGCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCCT
CAGCTTCCCCCGGCCCGGTGAGGCGGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGGCATCACACTGCCCTGTGCAGCGGAGCCGACCAGGCTCTTGTGTCTCA
CTCAGGTTTGCTTCCCTGTGCCCACTGCTGTATGATCTGGGGCCACCACCTGTGCGCGT
GGCCTCTGGGCTGCCCTCCCGTGGTGTGAGGGCGGGGCTGGTGTCTATGGCACTTCTCCTTG
CTCCACCCCTGGCAGCAGGGAAGGGCTTTCCTGACAACACCCAGCTTATGTAAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCAGACTCTGTC
TGTGCCGAGTGTATTATAAAATCGTGGGGGAGATGCCCGGCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

<subunit 1 of 1, 224 aa, 1 stop

MESGAYGAAGAGGSFDLRRFLTQPVVARAVCLVFALIVFSCIYEGEGYNAHESKQMYCVFN
RNEADACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLIVIGDLLFSALWTLFWVFGFC
FLTNTQWAVTNPKDVLVGADSVRAAITFSFSSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTF
DPNTAYASYPGASVDNYQOPPTQNAETTEBGYPPPVY

Type II Transmembrane domain:

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

GAGCCACCTACCTGCTCCGAGGCCAGGCTCGAGGGCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGGATG
 CCGCTGGCCGAGGCCCGCCAGGCTGGCTGGCGGGCAGGGGGACGGAGTGATGGCGAGGAAGCGGAGCCAGAGGGG
 ATGTTCCAGGCGCTGTGAGGACTCCCAAGAGAAAAGCCCGGGCTACCTCCGCTGGTGCCCTGTGTTGCTGCTG
 GCCCTGCTCGTGTGGCTTCGGCGGGGTGCTACTCTGCTATTTCCTAGGGTACAAGCGGGAGGTGATGGTCAGC
 CAGGTGTACTCAGGCAGTGTGCTGTACTCAATCGCCACTTCTCCAGGATCTTACCCGCGGGGAATCTAGTGCC
 TTCCGTCAGTGAACCCGCCAAGGCCAGAGATGCTCAAGGAGCTCATCACCAGCACCCGCTGGGAATCTACTAC
 AACTCCAGCTCCGCTCTATTCTTTGGGGAGGGACCCCTCACTGCTTCTTCTGTTTCAATTTCTCCAAATCCCCGAG
 CACCGCCGGCTGATGCTGAGCCCGAGGTTGGTGAGGCACTGCTGGTGAGGAGCTGCTGTCCACAGTCAACAGC
 TCGGCTGCCGTCCCCACAAGGCCGAGTACGAAGTGGACCCCGAGGGCCCTAGTGATCTCTGGAAGCCAGTGTGA
 GACATAGCTGCATTGAATTCCAAGCTGGGTGTTACCGCTACAGCTACGCTGGGCGAGGGCCAGGTCTCTCGGGCT
 AAGGGGCTGACCACTGGCCTCCAGCTGCTGTGGCACTGCAAGGGCCCAAGAGCTCTCATGCTCAAATCTCCGG
 CTGGAGTGGACGCTGGCAGAGTGCAGGACCGACTGGCCATGTATGACGTGGCCGGGCCCTGGAGAGAGGCTC
 ATCACTCGGTGTACGGCTGACAGCCGCCAGGAGCCCGTGGTGAGAGTTCTGGCGTCGGGGGCCCATCATGGCGGT
 GTCTGGAAGAAGGGCTGCACAGCTACTACGACCCCTTCTGCTCTCCGTGACGCGGTGGTCTTCCAGGCCCTGT
 GAAGTGAACCTGACGCTGGACAACAGGCTCGACTCCAGGGCGTCTCAGACCCCGTACTTCCCAGCTACTAC
 TATGCACTGAGGAGGAGAGAATGATGATTTCGCGTGCAACCAGGGCCAGTGAGAGATCCAGAACAGGAGGCTGTGT
 GGCTTGCGCATCTGTCAGCCCTACGCCGAGAGGATCCCGTGGTGGCCAGCGCCGGGATCACCATCAACTCAACC
 TCCCAGATCTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTTGACAACAGTCGAGACCCCTGCCCTGGA
 GAGTTCTCTGTTCTGTGAATGGACTCTGTGCTCTGCTGTGATGGGGTCAAGAGCTGCCCAACCGCCCTGGAT
 GAGAGAACTGCGTTGTGAGAGCCACATTCCAGTGCAAGAGGAGCAGCAGATGCATCTCACTGCCAAGGTCTGT
 GATGGGACGCTGATTGCTCAACCGCAGCGATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTCACC
 TTCCAGTGTGAGGACCGGAGTGCCTGGAAGAACCCCAACCGCAGTGTGATGGGCGGCCGATGCAAGGACCGG
 TCGGATGAGGAGCACTGTGACTGTGGCCCTCAGGGCCCTCCAGCGCATTTGTTGGTGGAGCTGTGTCTCCGAG
 GGTGAGTGGCCATGGCCAGCCAGCTCCAGGTTCCGGGTGACACATCTGTGGGGGGCCCTCATGCTGACCGC
 TGGGTGATAACAGCTGCCACTGCTTCCAGGAGGACAGCATGGCTCCAGCGCATTTGTTGGTGGAGCTGTGTCTCCGAG
 GGTGAGTGGCAGAACTCGCGCTGGCTGGGAGAGGTGCTCTCAAGGTGAGCGCGCTGTCTTGCACCCGTAACCA
 GAAGAGGACAGCCATGACTACAGCTGGCGCTGTGTCAGCTCGACACCCGCTGGTGTGCTGGGCGCGTGGC
 CCGCTGTGCTGCGCGCGCGCTGCCACTTCTTCGAGCCCGGCTGCATGCTGGATACGGGCTGGGGCGCTTG
 CGGAGGGCGGCCCATCAGCAAGCTCTGCAGAAAGTGATGTGCAGTTGATCCCAAGGACCTGTGACAGGCTG
 GCTATGCTACAGGTGACGCAAGCATGCTGTGTGCGCGCTACCGCAAGGCAAGAGGATGCTGTGACGGT
 GACTCAGGTGCTCGCTGTGTGTGCAAGGCACTCAGTGGCGCTGCTCTCTGGCGGGCTGGTCACTGGGGCTG
 GGCTGTGGCGGCCCTAACTACTTCCGGCTACACCCGCTACACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
 ACCTGAGGAAGTCCCCCTGCAAGAGCAGGGCCCACTCTGTGACTCAGAGGCCCGAGGCACTGCCAAGCAGG
 GGGACAAGTATTCTGGCGGGGGTGGGGAGAGAGCAGGCCCTGTGGTGGCAGGAGGTGCATCTTCTCTGCTCC
 CTGATGTGTGCTCCAGTGTATGGCAGGAGGATGGAGAAGTGCCAGCAGCTGGGGTCAAGAGCTCCCTGAGGAC
 CAGGCCCAACCCAGGCCCTTGTGCTCCCAATTCTCTCTCTCCGTCCCCCTCTCTCACTGCTGCTATGCAAG
 CGAGTGGCTCAGCAGCAAGAAATGCTGGTTCTACATCCCGAGGAGTGTCTGAGGTGCGGCCCACTCTGTACAGAG
 CTGTTTGGGCGAGCTTGCCTCCAGAGAGCAGATTCCAGCTTCGGAAGCCCTGGTCTTAATTTGGGATCTGGGAAT
 GGAAGTGTCTCCATCCGAGGGGACCCCTCAGAGCCCTGGAGACTGCAAGGTGGGCTGCTGCCACTGTGAAGCCAA
 AAGGTGGGAAGTCTGACTCAGGCTCAGGTCCTTGGCCACCCCTGCTGCCACTCGGGCTCAGAGCCCAAGCCCT
 CACTGGGAGGTGAGCTCAGCTGCCCTTGGAAATAAGCTGCTGATCAAAAAAAAAAAAAAAAAAAAA

10015177.102504

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152
><subunit 1 of 1, 802 aa, 1 stop
><MW: 88846, pI: 6.41, NX(S/T): 7
MPVAEAPQVAGGQDGGDGEAEPEGMFKACEDSKRKARGYLRVLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSGVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT
YNNSSSVYSFGEGPLTCFFWFILQIPEHRRMLSPVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGVRLKGPDLASSCLWHLQGPDLML
KLRLEWTLAECDRLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVVFQACEVNLTLNRLDSQGVLTSPYFSPSYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPTQGGWTIQNRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFCLSVNGLCVPACDGVKDCPNGLDERNCVCRAFPQCKEDSTCIS
LPKVCDDGQPDCLNGSDDEEQCQEGVPCGTFTFQCEDRSCVKKPNPQCDGRPD CRDGSDEEHCD
CGLQGPPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIA DRWVITA AHCFQEDSMASVTVL
WTVFLGKVMQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALLQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKQVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDGSGGPLVCKALSGRWFLAGLVSWGLGCGRPNFYGVYTRITGVISWIQQVVT

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

100161770001

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACTCCAGAT
CTCCCCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAACTGCGTTTGAGAGCCACATTCCAGTGC AAAGAGGA
CAGCACATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGAGTGTGATGGGCGGCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCCAGCCGATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGTGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT
GACTACGACGTGGCGTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCGCCGTGCGCCC
CGTCTGCTGCCCCGCGCTCCCACTTCTTCGAGCCCGGCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCACAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCTGGCGGGGCTGGTCAGCTGGGGCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCAGAGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAAGTGGCCCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

1001617.10501

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGG
TCCGTGCGCGCAAGTTTTCATTTTCCACCTTCTCTGCTCCAGTCCCCAGCCCTGGCCG
AGAGAAGGGTCTTACCGGCCGGGATTGCTGGAACACCAAGAGGTGGTTTTTGTTTTTTAA
ACTTCTGTTTCTTGGGAGGGGGTGTGGCGGGGCAGGATGAGCAACTCCGTTCTCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTTGCTGCGGGAGCCCCGTACCTTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTTAACCTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTCCGCCACAG
CCAGCCCTTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTTCATCATTACGGAT
GGACGATGAGCGGTATCTTTGAAAACCTGGCTGCACAACTCGTGTGAGCCCTGCACACAAGA
GAGAAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGA
TCCGGTCAATAATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGG
AGAAGGACGATTTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG
GCCGGGTATGCAGGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGC
CGGGCCCATGTTTGAAGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTG
TGGATGTCTCCACACCTACACGCGTTCTCTCGGCTTGAGCATTGGTATTCAGATGCCTGTG
GGCCACATTGACATCTACCCCAATGGGGGTGACTTCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAATGTGAGCATGAGCGAGCCG
TCCACCTCTTTGTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTTGCTTCCAGTGCACT
GACTCCAATCGCTTCAAAAAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAATGTACCTAAAAACCCGGG
CAGGCATGCCTTTAGAGGTAACCTTCAGTCCCTGGAGTGTCCCTGAGGAAGGCCCTTAATA
CCTCCTTCTTAATAACCATGCTGCAGAGCAGGGCACATCCTAGCCAGGAGAAGTGGCCAGCA
CAATCCAATCAAATCGTTGCAATCAGATTACACTGTGCATGTCTAGGAAAGGGAATCTTT
ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646
><subunit 1 of 1, 354 aa, 1 stop
><MW: 39362, pI: 8.35, NX(S/T): 2
MSNSVPLLCFWSLCYCFAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRYSKDPPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHLKLVSAHREKLDANVVVDWL
PLAHQLYTDVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLVDKPSFAFQCTDSNRFFKKGICLS
CRKNRCNSIGYNAKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

100517-10504

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPRIIFPNHACEDPPAVLLEVQGTLRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKHLACGSERLTRSPLQPLISLCEAPSPQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEETFQCLNHRCSVSAVQRCDGVDACGDGSDGAGCSDPFPGLTPRP
VPSLPCNVLTLEDYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRRLAVRFTALDLGFGDAVH
VYDGPGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVS YHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLSGSLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCDEKCVYETWVCDGQPD CADGSDWDSCS
YVLPKRKIVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMPGGGPGARRRQRGRIMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLP SLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA
GACGCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC
AAGCCCCTGAACCATATATTGTTTACTGGAATTTGAAGTCACCGTTATCTTATTTTTTCATA
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATAT
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACCTTGTGACAGCAGTATGCTGTCTT
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCAGCGGTCCTTACCAGAAAAA
GCCTGTGCATGAAAAAAAAAGAAGTTTGTTAATTTTATATTACTTTTGTGTTGATACTAAGT
ATTAAACATATTTCTGTATCTTCAAAAAAAAAAAAAAAAAAAAA

10501
10502
10503
10504
10505
10506
10507
10508
10509
10510
10511
10512
10513
10514
10515
10516
10517
10518
10519
10520
10521
10522
10523
10524
10525
10526
10527
10528
10529
10530
10531
10532
10533
10534
10535
10536
10537
10538
10539
10540
10541
10542
10543
10544
10545
10546
10547
10548
10549
10550
10551
10552
10553
10554
10555
10556
10557
10558
10559
10560
10561
10562
10563
10564
10565
10566
10567
10568
10569
10570
10571
10572
10573
10574
10575
10576
10577
10578
10579
10580
10581
10582
10583
10584
10585
10586
10587
10588
10589
10590
10591
10592
10593
10594
10595
10596
10597
10598
10599
10600

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645
><subunit 1 of 1, 152 aa, 1 stop
><MW: 17170, pI: 9.62, NX(S/T): 1
MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRDLRIMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

GGGCGAGAAGTAGGGGAGGGCGTGTTCCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAAACC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGTGCTGCTGGGTCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAAATAAACATCGCCCTTCTGCTTCAGTGTGAAGGCCACGTG
AAGATGCTCGGGTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAGCCCC
TGAACCATATATTGTTATCACTGGATTGTAAGTACCGTTATCTTATTTTTTCATACCTTTTAT
ATGTCATCAGACTTGATCGATTGAATGAAGTGGTTATTTTTGGCCCTTTGCTGATATTATCAAC
TCACTGGTAAACAACAGTATTCACTGATCATGTCATCTGTGTTGGCAGTCATACAGAAACAC
AACATTGCAGTTGGTGGAGGGGTGTTTCCTGCTGTGACAGCAGTATGCTGTNTTGGCGAC

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGAGCGGGACGGGACCGGGCTAGG
CTGGGCGCGCCCCCGGGCCCCGCCGTGGGCATGGGCGCACTGGCCCGGGCGCTGCTGTGTC
CTCTGCTGGCCAGTGECTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG
CCCCTCCGGGTGGCCGCGGCACGAAACCGCGTAGTTGCGGCCACCCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCGCGG
GCGCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACAGTACTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTGTCAA
CATTGCCACTATTTTGAATCAGAGAATTTCTTTTGGCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGTTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAATTTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCCTGTATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGCTGGACGAATTCGGAACACCTTGGTCTTACTTCCCTAAAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCAATCCGTATCACAATCCTGCCTCAGCTTTACATTTCAGC
CCATGATGGGGGCGGCTGAATTATGAATGTTACCGATTCCGGCATTTCCCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCAGAA
GAGGGTGGGCTTCGCAGCGAGCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGAATGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCTGCTGCTGCTGCCGTTCCGGTGTGAGCGTCGCCCCGTGACCCTGAGGTCGTCA
ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAATTGAATAGCCAGGCCCTGACCTCAAGCAA
CCATGAACCTCAGCTATTAAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCACCCGCTTTCAATCTCTGTCTCTCCAGATGCCTTCTAGATTAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSDGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVITPKGFNTSFLVNIATIFESENF
FLPGIKWNGILGLAYATLAKPSSSLETFDLSVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDWYTPPIKEWYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLR
LPQKVFDVAVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISIVLRDENSSRSFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLPFR
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCAGGCCCTGTGCAGGGCGGATCG
GCAGCCGCTTGGCGGCATCCAGGCGGTGCGGGGCTGGGCGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCCGCGGTGCGGCGGCATGGGCAACCTGCGGGGCCACAGGCCGTGG
TCACGGGCGCCAAAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGAGCG
CGCGTGGTGTGGCCTGCCGACGCCAGGAGCGCGGGGAGGCGGCTGCCCTTCGACCTCCGCCA
GGAGAGTGGGAACAATGAGGTCATCTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCTCATCCACAATGCC
GGTATCAGTTCTGTGGCCGACCCGTGAGGCGTTAACTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCCTGAAGGCATGTGCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCACTGTGCGGGACGTCTTGACTTCAAACGCCCTGGACCGC
CCAGTGGTGGGCTGGGCGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGCCCGGAGCTCGCCAACAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCACCC
CAGGGCCTGTGAACCTCGAGCTGTTCCCTGCGCCATGTTCTCGGATGGCTGCGCCCACTTTTG
CGCCCATTTGGCTTGGCTGGTGTCTCCGGGCACCAAGAGGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG
AGGTGCCCTCCAGCTGCCCCGAGACGACCGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCTGGGAGGATGCTGAACCGATGAAGACCCCACTGTGAGGACTC
AGAGGCCCATCTTCTCTAAGCACCCCCACCTGAGGAGCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTGTCTAAGATGACGCACCGAATTGAGGCTAAGATTGAG
CCTGAGATCCAGCTCTCCTTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCCTT
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCTTATTGATTCTG
ATCAGCTCTGAGCAGAGGCAGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAACCTGATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCCATGTTAATGAAGCG
GAATTAGGCTCCGAGCTAAGGGACTCGCCTAGGGTCTCAGTGAAGTAGGAGGAGGGCCTG
GGATCTGAACCCCAAGGCTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCATGGGAGTAAGGGGACGCGCTTCCGGGCGG
ATGCAGGGCTGGGTCTCTGTATCTGAAGCCCTCGGAATAAGCGCGTTGACCGCCAAAA
AAAAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVTGANSIGIKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLLRVNHIGPFLLTHLLLPCLKACAPSRVVVVASAAHCRGRLDKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSELFRLRHVPGWLRPLLR
PLAWLVLRAPRGGAAQTPLYCALQEGIEPLSGRYFANCHVEEVPFAARDRAAHLWEASKRL
AGLGPGEDAEPEDEDPOSEDSEAPSSSLSTPHPEEPTVSPQYPSPQSSPDL SKMTHRIQAKVEP
EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76

GGAGGAGACAGCCTCTGGGGGGCAGGGGTTCCTGCCCTGCTGCTCTGCTCATCATGGGAGGCATGGCTCAG
 GACTCCCGCCCCAGATCTAGTTCACCCCCCAGGACCAAGCTGTTCCAGGGCCCTGGCCCTGCCAGGATGAGCTGC
 CAAGCCCTCAGCCCGCAGCCACTCCCAACCATCCGCTGGTTGCTGAATGGGCAGCCCTGAGCATGGTGCCTCCCAAGAC
 CCACACCAAGCTCCTGCTGATGGGACCTTCTGCTGCTACAGCCCCCTGCCCGGGACATGCCACGATGGCCAG
 GCGCTGTCCACAGACCTGGGTGTCTACACATGCTGAGGCCAGCAACCGGCTTGGCACGGCCAGTGGACAGGCGCT
 CGGCTGTCTGGCTGTCTCTCCGGGGAGGATTTCCAGATCCAGCCTCGGGAATCTGGTGGCTGTGGTGGGTGAGCAG
 TTTACTCTGGAATGTGGGCGCCCTGGGSCACCCAGAGCCCAAGTCTCAAGTGGGAAGATGGGAAACCCCTG
 GCCCTCAGCCCGGAGGCACACAGTGTCCGGGGGTCCCTGCTGATGGCAAGAGCAGAGAAGATGACGAGGGG
 ACCTACATGTGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCCGCGCAGCCCGGGTTTCATCCAGGAGCC
 CAGGACTACACCGAGCTGTGGAGCTTCTGGCTGTGCGAATTACAGCTGGAAATGTGACACTGCTGAACCCGGAT
 CCTGCAGAGGGCCCCAAGCCTAGACCGCGGTGTGGCTCAGCTGGAAGGTCAAGTGGCCCTGCTGGCCCTGCCAA
 TCTTACACGGCTTGTTCAGGACCCAGACTGCCCGGAGGCCAGGGAGCTCCGTTGGGCAGAGGAGCTGCTGGCC
 GGCTGGCAGAGCGCAGAGCTTGGAGGCCCTCACTGGGGCCAAGACTACGAGTTCAAGGTGAGACCATCTCTGGC
 CGGGCTCGAGGCCCTGACAGCAACGTGCTGCTCTGAGGCTGCCGAAAAAGTGCCAGTGGCCCACTCAGGAA
 GTGACTCTAAGGCTGGCAATGGCACTGTCTTGTGAGCTGGGTCCCAACCTGCTGAAAAACCAATGGCATC
 ATCCGTGGCTACCAAGTCTGGAGCCTGGGCAACACATCACTGCCAACAGCCAACTGGACTGTAGTTGGTGAGCAG
 ACCCAGCTGGAAATGCCACCAATATGCCAGGCTCCTACTGCGTGCAAGTGGCTGCAGTCACTGGTGTGGAGCT
 GGGGAGCTGAGCTAGACTGTCTGCTCTTTTGGAGCAGGCCATGGAGCGAGCCACCCAGAACCCAGTGAAGCAT
 GGTCTCTGGACCTGGAGCAGCTGAGGGTACCTTGAAGCGGCTGAGGTCACTGGCCACTGCGGTGTGTCACCT
 TGGCTGTGCTCTGGGACCGCGGTGTATTCACCGCCGGCGCGAGCTAGGTGACCTGGCCCGCAGSTCTG
 TACAGATTAACCACTGAGGATGCCATCTAAAAACAAGGATGGATCAAGTGAATCCAGTGGTTGGCAGACACT
 TGGCGTTCCACTCTGGCTCTCGGGGACCTGAGCAGCAGCAGCCTCAGCAGTCCGCTGGGGCGGATGCCCG
 GACCCACTAGACTGTGTCGTCTCTGCTCTCTGGGACTGCCAAGCCCGCGGTGCCCTGCTTCCAGACA
 AGCATCTTTATGTGCTCTCATGCTGAGCTGCCCTCCAGTACCCAGCCAGGCCAAGTCCCAAGTCCAGCT
 CTCAGGCCCTCCCAACCCCACTGGCCAGCTCTCCAGCCCTGTTCAGCTCAGACAGCTCTCGAGCCGAGG
 GACTCTCTCTCCCGCTTGTCTCTGGCCCCTGACAGGCTTGGAGGCCAAAAAGCAGGAGCTGCAGCAT
 GCCACAGTTTCCCACTGCTCCGGGGCAGCCACTCTTGGAGCTCCGGGCTGTGAGTTAGGAAATAGAGGTTCC
 AAGACCTTTCCCAAAGCCAGGAGCTGTGCCCAAGCTCTGGTTGCTGGCGGGCTGGGACCGAACTCTCT
 AGTCAACAGACCCAGCTCTCCGTGGCCACCAAGGCTCCCTCTCATCTGCTGCACAGAGCCCATCCCATC
 CTTAGCCCTGCACTGCCCTAGCCCCCAGGCTCTTCCCTCTCTGGCCCCAGCCAGCTTCCAGTCGCTGTCC
 AGCTCTCTCACTGTCTCACTGGGGGAGATCAAGACAGCTGCTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
 CTCAGTGAGGGTGAGGAGATCCCAAGAACAGCGTCTCTCCATGCCAAGGGCTCTTCCACCCCAACCACTAT
 GGGTACATCAGCGTCCCAACAGCTCAGAGTTCAAGGACATGGGAGGTGAGCCACCCAGCGGCTGGGAGGGG
 GGAGTCTTGCTGTGGCCCTGCTCTCCGCTGCTCAACCCCAAGCCAGCGGAGCTCCTTAGCCAAATGGTTGGGC
 TCAGCTCTCAGGAGCAATGCCCGAGCGCCAGAGCCAGCTTGTGAGCTCTCCGATGGCTCTTCTGCTGTGAT
 GTCACCTTGGCCCGGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTTGGTCTAGAGCCAGGAGGCGAGACTGC
 GTCTCTATAGATGCTCATCACTTCTCTCCACGCGGATGAGATCTTCTGACCCCAACTCTCTCCCTGCCCTG
 TGGGAGTGGAGGCCAGACTGGTTGGAAGACATGGAGGTGAGCCACCCAGCGGCTGGGAGGGGAGGAGTGCCTCC
 TGGCCCCCTGACTCTCAGATCTCTTCCAGAGAAAGTCAAGTCCACTGTGATGGCCCAAGCTGTGCTTCTCT
 GTAGATTACTCTGAAGAGCTGCTTCCAGAGTTCACAGAGGGGAATCAGAAACATCTCTCTGTCCACCAAG
 ACCTGGGCTGTGGTGTGGGTCTTGGCCCTGTGTTTCTTGCAGCTGGGGTGCACCTTCCCAAGCTCTCCAGAGG
 TTTCTCCCTCAGAGCTTGTGAAAACAAATGAAACAAAAATAGAGCAAGCTGACCTGGAGGCTCAGGGAGCAAA
 ACATCATCTCACTCACTGCTCTAGCCACTGCTTCTCTCTGTGCCATCACTCCCAACCAAGGTGTGTTTGGC
 CTGAGGAGCAGCCTGACTGTGCTCTTCCCAACCATTTGGATCAGAGGAAGTGGAGAGGAGGAGGCTGCTT
 GTGGAGGACAGCAGTGGCTGTGGAGAGGGCTGTGGAGGAAGGAGCTTCTGGAGCCCTCTCAGCTCTTACT
 GGGCCCTCTCTCAGAGAAGAGCTCAACTCTCTCCCACTCACTGAAAGAAATTAATTATGAATGCCACTG
 AGGCACTGAGGCCCTACTCTATGCGCAACCAAGGGTCAAGGCTGGGTCTAGAGGAGGATGCTGAAGGAAGGAGG
 TATGAGACGTAGGTCAAGAGCACCATCTCTGATCTGTGTGCTCACTATGAGCTTGAAGAAATTTGATACCAATAAAT
 GGTAAAAA

1016177-102504

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404
<subunit 1 of 1, 985 aa, 1 stop
<MW: 105336, pI: 6.55, NX(S/T): 7
MGGMAQDSPPQILVHPQDQLFQGPGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHLLLP
DGTLLLLQPPARGHAHDGQALSTD LGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAEKSDGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENVTL LNPDPAEGPKRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGFW
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYS EDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRS PGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCCSSSDSLCSRRLSSPRLSLAPAEA
WKAKKKQELQHANS SP LLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPAPLPHETPPTQSQQTPPVAPQAPSSILLPAAPIILSPCSPSPQASS
LSGSPASSRLSSSSSLSLGEDQDSVLTPEEVALCLBLSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGR TG GVGPKGGVLLCPPRPCLTPTPSEGLANGWGSASEDNAA SARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLPEPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSH TQRLGRGMPPWP PDSQISSQRSQLHCRMPKAGASPV DYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCTCAGCGCCAGAAATGCGGCTTCTGGTCTGTATGGGGTGCCTGCTGCT
 CCCAGGTTATGAAGCCTTGAGGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
 CCCTGTCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
 GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGGCCAGGAGACAAT
 GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCTGTGGA
 ACCTCACCTTGAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
 TCTTTACTGATCTCTCTGTTGCTCTTCCAGGACCTGCTGTCTCCCTCCCTCTCTCCAC
 CTTCCAGCCTCTGGCTACAACACGCTGACGCCAAGGCAAAAGCTCAGCAAAACCCAGCCCC
 CAGGATTGACTTCTCTGGGCTCTACCCGGCAGCCACCACAGCCAGGAGGGGAAGACAGGG
 GCTGAGGCCCTCCATTGGCAGGGACTTCCAGTACGGGCAGAAAGGACTTCTCAGTACAC
 AGGAACCTCTCCTCACCCAGGACCTCTCCTCTGAGGGAGCTCCCGCCCCCCCCATGCAGC
 TGGACTCCACCTCAGCAGAGGACACAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
 GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTCTGGTGTCTGTGAGCCTTCTGTGAGC
 CGCAGGCCGTGATCGCCTTCTGCGACCACTGCTCTCTGTGGAGAAAGGAAGCTCAACAGGCCA
 CGGAGCACACAGGGAACGAGAAGTTCTGGCTCTCAGCCTTGACTGCGGAGGAAAGGAAGCC
 CCTTCCCAGGCCCTGAGGGGGACGTGATCTCGATGCTCTCCCTCCACACATCTGAGGAGGA
 GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCTGGCCAGGCCAGCAGT
 GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCGAAAGCTTTCCACCTCAGCCTCAGAG
 TCCAGCTGCCCGGACTCAGGGCTCTCCCCACCTCCCCAGGCTCTCTCTTGCATGTTTCCA
 GCCTGACCTTAGAAGCCTTTGTGAGCCTTGGAGCCAGAGCGTGGCCTTGTCTTCCGGCTG
 GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCGAGTACCAGGCTGCTGACCCTCA
 GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGC
 CTCATGCCAGTGTGCGAACCTGCCTTCTCCCACTCCAGACCCACCTTGTCTTCCCTCCC
 TGGCGTCTCAGACTTAGTCCACGGTCTCCTGCATCAGCTGGTGTGATGAAGAGGAGCATGCT
 GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTCAGGAAGCCT
 GTGAAAAACGTGATTCTTGGCCCCACCAAGACCACCAAAACCATCTCTGGGCTTGGTGCAG
 GACTCTGAATTTAAACAATGCCAGTGACTGTGCGACTTGAGTTTGGAGGCCAGTGGGCCCTG
 ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCACTTGGGCTGTGACGCTCTCCACCTGCC
 CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGGACTCCCCTGAGGCCTGCTAAG
 TCCAGGCCCTTGGTCAGGTCAAGTGCACATTGCAGGATAAGCCAGGACCGGCACAGAAGTGG
 TTGCCCTTINCCATTGGCCCTCCTGGNCCATGCCTTCTTGCTTTGGGAAAAAATGATGAAGA
 AAACCTTGGCTCCTTCTGTCTGGAAGGGTTACTTGCCCTATGGGTCTCTGGTGGCTAGAGA
 GAAAAGTAGAAAAACAGAGTGACAGTACGTAGTGTCTAACACAGAGGAGAGTAGGAACAGGGCCG
 ATACCTGAAGGTGACTCCAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA
 GCACACTACTATTCTTTTCTTTTCTCATTATATTGTTTTTAAAGACAGAATCTCGTGCT
 GCTGCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCCTCTGGGTTCAAGTGATT
 CTTCTGCCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACACCTGGGTAAAT
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCTGAC
 CTCAAATGAGCCTCTGCTTCTCAGTCTCCCAAATGGCCGGATTACAGGATGAGCCACTGTG
 TCTGGCCCTATTCTCTTAAAGAGTAAATTAAGAGTTGTTGATGATCAAAACTTGGAAAG
 ATGGAGGAGAAAAAGGAAAGAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
 TATTTCTGTTTTTGTGATCTCCTTCCACTCTTTTCTTCTTACATAAATTTGCCGGTGTCTTT
 TTTACAGAGCAATATCTGTGATATACAACCTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
 ATCATTATTCTCAGCACTTCTCTGTGTTTTACAGACCTTTTATAAAATAAATGTTATCATCA
 GCTGCATAAAAAAAAAAAAAA

100577-10504

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAEEEGQETMKGRVSIIRDSRQELSLIVTLWNLTLDQAGEYWCQVEKRGPDSELLISLFV
FPGPCCPSPSPSTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGTGAEAPPLPG
TSQYGHERTSQYTGTSPPHATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAAEKEAPSQAPEGD
VISMPPLHTSEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAGAACCAGTGGTCTC
GCTCTGTTGCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACACGAGAGGTGGTGAGCTGCCTGTTCATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTCACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCAGATCAACAGCCGGAGGTGGTGCAGCAACCTACCCCGAACGTCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGTTTCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCNSLTPNVPNVCRMVCSDLLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHCQGKDLTEWVDGCDP

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCGGGCCGGGCGCCCGCGGCGGCACCATGAGTCCC CGCTCGTGCCCTGCTTC
 GCTGCGCCTCCTCGTCTTCGCGCTCTTCTCAGCCGCGCGAGCAACTGGCTGTACCTGGCCA
 AGCTGTCTGTCGGTGGGAGCATCTCAGAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATC
 CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTTCATGGACTCGGTGCGCCGCGTGC
 CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAACCTGCTCCACAC
 TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGACTCGGGAGGCGGCCTTCGTG
 TACGCCATCTCTTCGCGAGGTGTGGCCTTTGCAGTGACGCGGCGTGACAGTGGGGAGCT
 GGAGAAGTGCGGTGTGACAGGACAGTGCATGGGGTCAGCCCACAGGGCTTCCAGTGGTCAG
 GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG
 AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACATGAGGCCGGCAG
 GAAGGCCATCCTGACACACATGCGGGTGAATGCAAGTGCCACGGGGTGT CAGGCTCCTGTG
 AGGTAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAG
 AAGTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACC
 ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGAGCCTAGCCCCG
 ACTTCTGTGAGCAGGACATGCGCAGCGCGTGTCTGGGCACGAGGGGCCGCACATGCAACAAG
 ACGTCCAAGGCCATCGACGGCTGTGAGCTGTGTGCTGTGGCCGCGGCTTCCACAGGCGCA
 GGTGGAGCTGGCTGAACGCTGCAGCTGCAAAATCCACTGGTGCTGCTTCGTCAAGTGCCGCG
 AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATGACCGCCTGCGCTAGCCCTGCGCCGCG
 AACCACTAGTGGCCAGGGAAGGCGATAATTTAAACAGTCTCCACCACCTACCCCAAGA
 GATACTGGTTGTATTTTTTGTCTTGGTTTGGTTTGGGTCTCATGTTATTATTATTGCGGAA
 ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT
 GCCACTGACCAAAAGGACCTTGCTCGTGCCGCTGGCTGCCCGCATGTGGCTGCCACTGACCA
 CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGACAGACCTAAGGTGGAGTAACAAGGAGTAT
 TACCACCACATGGCTACTGACCGTGTATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA
 GGTACCGACTTGATGGAAGTCACCCCTCTGGA AAAAAGAACTCTTA ACTCTCCAGCACACA
 TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG
 GAACAAGCAGATACCAGGTCAAGGGCACCAAGGTTCAATTCAGCCCTTACATGGACAGCTAGA
 GGTTCGATATCTGTGGGTCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA
 GTCCACCCCTAGAACCAGCCTGCCCCAGCCTGCCCTGGGAAGAGGAAACTTAACCACTCC
 CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT
 TGCAGTCATGCCCCAGTCACTTTTACAGCGCTGTTCTCCTCATGAAACTGAAAAACACACAC
 ACACACACACACACACACACACACACACACACACCGACACACACACACACCTGCGAGA
 GAGAGGGAGGAAAGGGCTGTGCCTTTGCAGTCATGCCGAGTCACCTTTCACAGCACTGTTCCTC

FIGURE 83

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCLSRLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRWNCSTLDSLVPVFGKVVTQGTREAAFVYAISSAGVAFVAF
TRACSSGELEKCGCDRTVHGVSPQGFWSGCSDNIAGVAFSQSFFVDVRERSKGASSSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCCKFWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCAACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCCTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCTCTGGTC
ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTGCGAGACTGCCACAGCTGCTGCTCGGGGACGACGGCGCAGCTG
CAGACCA CGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAAC TGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACCTCTGCGAGCCG
TGCCCCACGTCGTGGCTGTCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGACCTGGTGATCGTTGGGGGCC
TGGATGAGCAGGGCTTCTCACTCGGAACACGCGTGCCGTGGCCGTGGTTACTGGCTGGGCCTGAGG
GCTGTGCGCCATCTGGGAAGGTTTCAGGGCTACCACTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAAC CAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAAGTGTGTCTATGATGC
TGACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAAC TGGACCCCGCCAGTGCCCTGGAGCCGCGCCATTGCAGCATGTCTGTA
TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTCTTCTCT
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC
TGGGCTCTGGGACCTCCATGCGGACCTCATCCTAACTCCACTCACGCAGACCCAACTAACC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCGTGATATGCCTCCACTTCTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACTGGA
AGCTGTTTTTGACGCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGGSSSEVPGGPWGRVHWSRRPLFLALAVLVTTLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVTTQGLAEAGRGREDVRTELFRALAEAVRLQNNSCPECPSTWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLEQGFLLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFSS
HWNQGEPNDAWGRENVCVMMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGCATCCGACCCGGGGAAGGTCGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCGCCCCCGCAGCCCCCTTCTCCTCCTTTCTCCACGTCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAAC TGGAGCCTCATTGGCCGGCCCGG
GGCGCCGGCCTCGGGCTTAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCCGTGCCGGC
CGCGCTCCGCTGCTCCTGCCGGGTGATGGAAAAACCCAGCCCGGCCGCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCTGGCCACTCTCGGCGCCGCGGCCAGCCTCTTGGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCCCTGTTCCGCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTAGTAACGGGCTGCG
CGACTTTGCGAGCGCGGCGAGGCTGGGCGCTGATGAAGGAGATCGAGGCGCGGGGGAGG
CGCTGCAGAGCGTGCACGAGGTGTTTTGCGCGCCCGCTCCCAAGCGGCACCGGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGTGGTGCGCATCGTGCC
CAGCCCCGACTGGTTCTGTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGCGCTGGACCTGTACCCCTACGACGCCGGGACGGACAGCGGCTTACCTTCTCC
TCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCTCCTCTCCAG
CCACCCGGCCAACCTCCTTCTACTACCCGCGGCTGAAGGCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCGCCCCAGTCTTGCCCAGC
AGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGAGTCGAGGTTCTC
CCTGTGGTCTCCTGGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGTACGTCCGGGTCCAGCCCGCCAACAACGGGAGCCCTGCCCGAGCTCGAAGAAGAG
GCTGAGTGCCTCCCTGATAACTGCGTCTAAGACCAGAGCCCCGCGAGCCCTGGGGCCCCCG
GAGCCATGGGGTGTGGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCGGAGGGCACAGGG
GGTTTTCGCGCTGCTCTGACCGCGGTGAGGCGCGCCGACCATTCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG
CCCCCGTGTCCTGCTGCTCTCAGCCTCCTCCTCTGAGGATAAAGTCATCCCAAGGCTC
CAGTACTCTAAATTATGTCTCCTTATAAGTTATTGTGCTCCAGGAGATTGTCCTTCATCG
TCCAGGGGCTGGCTCCACGTGGTTGCAGATACCTCAGACCTGGTGCTCTAGGCTGTGCTG
AGCCCACTCTCCGAGGGCGCATCCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAA

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAALGKALCALLLATLGAAGQPLGGESIC SARAPAKYSITFTGKWSQTAFKQYPL
FRPPAQWSSLLGAAHSSDYSMWRKNQYVSNGLRDF AERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDFVGVDSLDCGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLLRLRQSP
RAFIPPAVPLPSRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGR LGTKSRTRYVRVQPA
NNGSPCPELEEEAECPDNCV

Important features:

Signal peptide:

amino acids 1-26

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCGTCGCCGGCGAGAGAGATGTCGCCCGG
CCCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCAGACCCATTTGCGCTTGCTGACGGCGTCTG
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCGGGACTCTGGGCTCCACCACC
GTGGCCGCCGGCGGGACACGACAGGCCGGCGTTTCTCCTTCGGAACGGGAACGCTAGCAA
CCCTTCTGTGGGGCTCAATTTTGGAAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGGTTTTTGGAAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATA CAGGTGCCCTTGACACCAAGAGCCCTCAAGTGGTCAACAAATATGGAAC
CCTGCAAGGAAAA CAGATGCATGTGGGGAAGACCCCATCCAAGTCTTTTTAGGAGTCCCTT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGACCTCCAGAACCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACCACTACCGCCTGGATGGAGTCTCGCTCTGTGCGCAGGCTGGAG
TGCAGTGGCAGCATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCTGTC
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCTCGGGCCAGCTGGCCCTCG
ATGTACGTGACGACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCGCCGGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGTAGTGGTCTGGTTCC
CGGAGGGCGCTTCATCTGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCGCG
GAGAAAGTGGTGTCTGTGTTTCTGCAGCACAGGCTCGGCATCTTCGGCTTCTGTAGCACGGA
CGACAGCCACGCGCGCGGGAACCTGGGGCTGCTGGACAGAGTGGCGGCTCTGCGCTGGGTGC
AGGAGAACATCGACGCTTCGGGGGAGACCCAGGAATGTGACCTGTTTCGGCCAGTCCGGCG
GGGGCCATGAGCATCTCAGGACTGATGATGTACCCCTAGCCTCGGCTCTCTTCCATCGGGC
CATTTCCAGAGTGGCACC CGCTTATTCAGACTTTTCATCACTAGTAACCCATGAAAGTGG
CCAAGAAGTTGCCCACTGGCTGGATGCAACCAACAGCACACAGATCTCGGTAACCTGTC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCTCCAACT
GAACTTCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
TGATCCAGATGACCTTTGGTGTCTCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT
CTAGTGTCAACAACCTGGAATTCAATTGGCTCTTGCTTTATAATATCAACAAGGAGCAGGT
ACCACCTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTTCGTGTATGCCACATGCGAGACTGTCTCAC
TACCACCGAAGAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCA TGAAGCTCAAGGAGAAGAAGATGGC
TTTTTGGATGAGTCTGTACAGTCTCAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGGTTTGCCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
TGGACATACCTGGGGACAAGAGTTCTACCCACCCAGTTTGAAGCTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCCTGTTGTGTGGGACCTGCATGCGCTTTCCAGCC
TGACATCCCATGATGCCCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACCTGTGCTCAGCTCTCCAGCCTCAGGACAACTCTTTTTTCCCTTCTTCAAACTCT
CCACACTTCAATGTCTCCTTGTGACTCCTTCTATGGGAGGTGCAGCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTTGGCATTTACCATCCATCCTGCTCAACCTTGTCTCTGTCTGT
TCAATTTGGCCTGGAGGCTAGGGCAGGTTGTGACATGGAGCAAACTTTTGGTAGTTTGGGA
TCTTCTCTCCCAACCACTATCTCCCCCAGGGCCACTCAAAGTCTATACACAGGGGTGG
TCTCTTCAATAAAGAAGTGTGTATAGAAAAA

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179
<subunit 1 of 1, 545 aa, 1 stop
<MW: 58934, pI: 9.45, NX(S/T): 4
MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATSAPSSG
FGTGLFGSKPATGFTTLGGTNTGALHTKRPQVVTKYGTLLQKQMHVGKTPIQVFLGVPPFSRPP
LGILRFAPPEPEPEPWKGIRDATTYPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNIVYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTFLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLACGNHNSTQILVNCLRALSGTKVMRVSNMKMRFLQLNFORDPPEEIIWSMSPVVDGVVIPPDD
PLVLLTQGVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMMMD
IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 92

GAGAACAGGCCTGTCTCAGGCAGGCCCTGCGCCTCCTATGCGGAGATATGCTACTGCCACTGCCT
 GCTGTCTCGCTGCTGGGCGGGTCCCAGGCCTATGGATGGGAGATTCTGGATACGAGTGCAGG
 AGTCAGTGAATGTTGCCGAGGGGCTGTGCATCTCTGTGCCCTGCTCTTTCTCCTACCCCCGA
 CAAGACTGGACAGGGTCTACCCAGCTTATGGCTACTGGTTCAAAGCAGTGACTGAGACAAC
 CAAGGGTGCTCCTGTGGCCACAAACCACAGAGTGCAGAGGTGGAATAGACACCCGGGGCC
 GATTCCAGCTCACTGGGGATCCCCCAAGGGGAACCTGCTCCTTGGTGATCAGAGACGCGCAG
 ATGCAGGATGAGTCACAGTACTTCTTTCCGGTGGAGAGAGGAAGCTATGTGACATATAATTT
 CATGAACGATGGGTTCTTTCTAAAGTAAACAGTGCTCAGCTTCACGCCAGACCCAGGACC
 ACAACACCGACCTCACCTGCCATGTGGACTTCTCCAGAAAGGGTGTGAGCGCACAGAGACC
 GTCGACTCCCGTGTGGCTATGCCCCAGAGACCTTGTATCAGCAITTCAGCTGACAACAC
 GCCAGCCTGGAGCCCCAGCCCCAGGGAAATGTCCCATACCTGGAAGCCCCAAAAGGCCAGT
 TCCTGCGGCTCCTCTGTGCTGCTGACAGCCAGCCCCCTGCCACTGAGCTGGGTCCTGCAG
 AACAGAGTCTCTCCTCGTCCCATCCTTGGGGCCCTAGACCCCTGGGGCTGGAGCTGCCCGG
 GGTGAAGCTGGGGATTTCAGGGCGCTACACCTGCCGAGCGGAGAACAGGCTTGGCTCCCGC
 AGCGAGCCTCGGACCTCTCTGTGCAGTATCCTCCAGAGAACCTGAGAGTGAATGGTTCCCAA
 GCAAACAGACAGTCTTGGAAACCTTGGGAACGGCACGTCTCTCCAGTACTGGAGGGCCA
 AAGCCTGTGCCTGGTCTGTGTACACACAGCAGCCCCCAGCCAGGCTGAGCTGGACCCAGA
 GGGGACAGGTTCTGAGCCCTCCAGCCCTCAGACCCCGGGTCTTGGAGCTGCCTCGGGTT
 CAAGTGGAGCAGAAGGAGAGTTCACTGCCACGCTCGGCACCCAGTGGGCTCCGACGAGT
 CTCTCTCAGCCTCTCCGTGCCTAATAAGAGGGACTCATCTCAACGGCATTCTCCAACGGAG
 CGTTTCTGGGAATCGGCATCAGCGCTCTTCTTTCTCTGCTGGCCCTGATCATCATGAAG
 ATTCTACCGAAGAGACGAGTCTCAGACAGAAACCCCGAGGCCAGGTTCTCCCGGCACAGCAC
 GATCTTGGATTACATCAATGTGGTCCCGACGGCTGGCCCTTGGCTCAGAAGCGGAATCAGA
 AAGCCACACCAACAGTCTCGGACCCCTCCTCCACAGGTGCTCCTCCCGAGAATCAAAG
 AAGAACCAGAAAAGCAGTATCAGTTGCCAGTTTCCAGAACCCAAATCATCTCACTCAAGC
 CCCAGAATCCCGAGGAGCCAGAGGAGCTCCATTATGCCACGCTCAACTTCCAGGCGTCA
 GACCAGGCGCTGAGGCCCGGATGCCCAAGGGCACCCAGGCGGATTATGCAGAAGTCAAGTTC
 CAATGAGGGTCTCTTAGGCTTTAGGACTGGGACTTCGGCTAGGGAGGAAGGTAGAGTAAGAG
 GTTGAAGATAACAGAGTGCAAAGTTTCTTCTCTCCCTCTCTCTCTCTTCTCTCTCTCTCT
 CTCTCTTCTCTCTCTTTTAAAAAACATCTGGCCAGGGCACAGTGGCTCAGCCTGTAACT
 CCAGCACTTTGGGAGGTTGAGGTGGGCAGATCGCCTGAGGTCCGGAGTTCCAGACAGCCTG
 GCCAACTTGGTGAAACCCCGTCTCTACTAAAAATACAAAAATAGCTGGGCATGGTGGCAGG
 CGCCTGTAATCTTACTTGGGAAGCTGAGGCAGGAGAACTCACTTGAACCTGGGGAGACGG
 AGGTTGAGTGAGCCAGATCACACCATGACGCCAGCCTGGGCAACAAAGCGAGACTCCA
 TCTCAAAAAAAAATCTCCTAAATGGGTTGGGTGTCTGTAATCCAGCACTTTGGGAGGCTA
 AGGTGGGTGGATTGCTTGGAGCCAGGAGTTCGAGACAGCCTGGGCAACATGGTGAAACCCC
 ATCTCTACAAAAAATACAAAAATAGCTGGGCTTGGTGGTGTGTGCTGTAGTCCCAGCTGT
 CAGACATTTAAACCGAGCAACTCCATCTGGAATAGGAGCTGAATAAAATGAGGCTGAGACC
 TACTGGGCTGCAATTCTCAGACAGTGGAGGCACTTAAGTCAAGGATGAGACAGGAGGTCCG
 TACAAGATACAGGTCAATAAGACTTTGTGTATAAAACAGATTGCAGTAAGAGAGCCAAACCA
 ATCCACCAAAACCAAGTTGGCCACGAGAGTGACCTTGGTGTGCTCACTGTACTACTCTCT
 GACAGCACCATGACAGTTTACAAATGCCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
 AAAGGGGGAGGAATGAATAATCCACCCCTTGTTTAGCAAAATAGCAAGAAATTAACCATAAA
 TGGGCAACCCAGCTCAGGCGTGTCTTGTCTATGGAGTAGCCATTTCTTGTTCCTT
 TACTTTCTTAATAAATCTGCTTTACCTTAAAAAAA

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002
><subunit 1 of 1, 544 aa, 1 stop
><MW: 60268, pI: 9.53, NX(S/T): 3
MLLP LLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFYPRQDWTGSTPAYGYWFK
AVTETTKGAPVATNHQSRREVEMSTRGRFQLTGDPKAGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFFLKVTVLVSFTPRPDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHWPWGPRLP
GLELPVGKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVMSQANRTVLENLNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVVRPRPEARMPKGQTADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins signature.

amino acids 365-371

1006770364

FIGURE 94

TGAAGAGTAATAGTTGGGAATCAAAAAGAGTCAACGCAATGAACCTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCCTCTCCTATGGCCTTGTCTTGGAGCAACAGAAAACCTCTCAACAAAGA
AAGTCAAGCAGCCAGTGCATCTCATTGAGAGTGAAGCGTGGCTGGGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACCTCGGCAGCTAAGATCTGATTTAGA
CAATGGAAACAATTCCTTCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TCATCTCTTAAGAGCCAGGTTAATAGACATCGCTACTTGGAGGGCTGTGGAACTGTGAGTCTGA
GTTTGTGTCATCAAAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG
AGGCCATTGTACAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTCTACAGCTTACTCAAGGCCAGCC
ATATTTTCTGTTGAACCAACAACAGGAGTCATAAGAAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTTGGGTAACTATTCAAGCCAAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAAACAACAGTGTATTAATTTAACTTTTCTGAGTGTAAATGACAATAAGCCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTTGAAGAGGAT
GATTCCGCAACATTTGACATTATTACTAATCATGAACTCAAGAAGGAATAGTTATATTTAA
AAAGAAAGTGGATTTTGGAGCACCAGAACCCTACGGTATTAGAGCAAAAGTTTAAAAACCATC
ATGTTTCTGAGCGCTCATGAAGTACCACACTGAGGCTTCCACCACTTTTCAATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCCTCTTTTCCCTCTCCATATATTGTTATTGAAAGTTTTTGA
AGAAAACCCACAGGATCATTTGTAGGCGTGGTGTCTGCCACAGACCCAGACATAGGAAT
CTCCTATCAGGTATCTATTACTAGGAGCAAAGTGTTCAATATCAATGATAATGGTACAATC
ACTACAGTAACCTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAAATATAGAACAGATCTCTCGATCCCACTGTATGTGCAAGTTCTTAAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAAATGCAAGCTCT
GGTCAGGTAATTCAGACTATCAGTGCACTGGATAGAGATGAATCCATAGAAAGACCACTTT
TTACTTTAATCTATCTGTAGAAGACACTAACRAATTCAGTTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCAATTTGACTAATAGAACTGGTTTTTAACTTCAAGAAGAACTGTCTTC
TACATCTCCATCTTAAATGCGGACAAATGGAATCCCGTCACTTACAAGTACAAACACCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTGTCATTATGATCAT
TTTGGTTTTATTTTGTGACTTTGGGTTTAAACAAACGGAGAAAACAGATTCTATTTCTGA
GAAAAGTGAAGATTTTCAAGAGAAATATATCCAAATATGATGATGAAGGGGGTGGAGAAGAAG
ATACAGAGGCTTTGATATAGCAGAGCTGAGGAGTAGTACCAATACGCGGAACGCAAGACT
CGGAAAAACCAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATTCAGGAATTCATTCTGGAAAAGCTCGAAGAAGCTAATCTGATCCGTGTG
CCCCCTCTTTGATTTCCCTCCAGACCTACGCTTTTGGGGAACAGGGTCAATTGACTGGATCC
CTGAGCTCTCTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTTAATGAGTT
GGGACCTCGCTTTAAAGATTAGCATGCTGTTGGTTCTGCAAGTCAGTCAATAAATTTAGG
GCTTTTACCATCAAAATTTTTTAAAGTGCTAATGTGTATTGAAACCAATGGTAGTCTTAA
AGAGTTTTTGGCCCTGGCTCTATGGCGGGGAAGCCCTAGTCTATGGAGTTTCTGATTTCC
CTGGATAAATATCCATGGTGTATTTTAAAGTACCTACATGCTGTGATGAAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACACAGATTTGAAGTAAATTAATG
TAGGAAGATATTTAAAGTAGATGAGAGGACACAAGATGTAGTGCATCCTTATGCGATTATAT
CATTTATTACTTGAAGAGGTAAAAATACCAACGAGAAAATTTAAGGAGCAAAAATTTG
CAAGTCAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATTGAAA
ATTGAAATGTATAGTCAGAGAAATTTTCAATGAATATTCCATGAAGTATTGTTTCTTTAT
TTAA

10016177-102501

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGVWVNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFYQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSNDINDNEPKFLDEPYEAIVPMSPEGTLVIQVTASDADDPSSGNNARL
LYSLQQGPYPYFVEPTTGIVIRISSKMDRELQDEYVWIIQAKDMIGQP GALSGTTSVLIKLS
VNDNKPIFKESLYRLTVSESAPTGTSTIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGIVILKKKVD FEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTI SAVDR
DESI EEHHFYFNLSVEDTNNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVDCDGS GSTQTCTCYQELVLSMGFKTEVIIAILCIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENI FQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEBANTDPCAPPFDSLQTYAFEGTGLAGSLSSLES AVSDQD
ESYDYLNLGPRFKRLACMFGSAVQSNN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTCAAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGNTGGAACAACAAGTGATTAATTAAACTTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATGAAGAGGATGATTGCAAAACATTTGACATTATT

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCGGGCGCGGACCCCAACCCCGAC
 CCAGAGCTTCTCCAGCGCGCGCGCAGCGAGCAGGGCTCCCCGCCTTAACTTCTCCGCGGGG
 CCCAGCCACCTTTCGGGAGTCCGGGTTGCCACCTGCAAACTCTCCGCCTTCTGCACCTGGCA
 CCCCTGAGCGCAGCGCGGGCCCCCGAGCGAGTCA**TGGCCCAACCGCGGGCTGCAGCTGTTGGGC**
 TTCATTCTCGCCTTCTGGGATGGATCGGCGCCATCGTCAGCACTGCCTGCCCCAGTGGAG
 GATTTACTCCTATGCCCGCGACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGA
 TGTCTTCGCTGTCGAGAGCAACGGGCAGATCCAGTGCAAACTCTTTGACTCCTTGCTGAAT
 CTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTAT
 AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGGAAGCAGATGAGGTG
 AGAAGATGAGGATGGCTGTCTATTGGGGGTGCGATATTTCTTTCAGGCTGGCTATTTTA
 GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCTATGACCCAGT
 CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGTGCTTCTCTGCCC
 TTCTGGGAGGTGCCCTACTTTGCTGTTCTCTGTCCCCGAAAAACAACCTCTTACCCAAACCA
 AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTG**TGAC**CACAGAGGCAAAAG
 GAGAAAATCATGTTTGAACCAAACCGAAAATGGACATTGAGATACATCATTAACTAGGAC
 CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAACAACAACAACAAAA
 ACCCATGTGTTAAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTCTCA
 ATATAGGAGGGAAGATTTTCCATTGTATTACTGCTCCACTGAGTAATCATACTCATCTCAAT
 GGGGGAAGGGGTGCTCTTAAATATATATAGATATGTATATATACATGTTTTTCTATAAAA
 ATAGACAGTAAAAATCTATTCTCATTATGTTGATACTAGCACTTAAATATCTCTAAAT
 AGGTAATGTATTTAATTCATATTGATGAAGATGTTTATTGGTATATTTCTTTTTCTGCC
 TTATATACATATGTAACAGTCAAATATCATTTACTCTTCTCATTAGCTTTGGGTGCCCTTG
 CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTCAATTCTTCATGCTGGCCCTTT
 CATATACTTATTTATTTTTTACCATAATCTTATAGCACTGCATCGTTATTAAGCCCTTAT
 TTGTTTGTGTTTCATGGTCTCTATCTCCTGAATCTAACACATTTCATAGCTACATTTTA
 GTTCTTAAAGCCAAAGAAGATTTATTACAAATCAGAACTTTGGAGGCAAATCTTCTGCAATG
 ACCAAAGTGATAAATTCCTGTTGACCTTCCCACACAATCCCTGTACTCTGACCCATAGCACT
 CTTGTTTGTCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTTGT
 AACACAACCTTATTGATTGAATTTTAAAGCTACTTATTCATAGTTTTATATCCCCCTAAACT
 ACCTTTTGTTCCCCATTCCTAATTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTTA
 TATCTTCTTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
 ATCTGGTGCAAAATTTCTCTCTGTAGCTGTAAAGCAAGTCACTTAATCTTTCTACCTCTTT
 TTCTATCTGCAAAATTTGAGATAATGATACTTAACAGTTAGAAGGGTGTGTGAATATTAA
 TTAGTTTATATTACTCTTATTCTTTGAACATGAACACTATGCCTATGTAGTGTCTTTATTGTCT
 CAGCTGGCTGAGACATGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
 CACTGCTTCTCTCTTACCAGTCTATTTCCACTGAACAAAACCTACACACATACCTCTCAT
 GTGGTTCACTGTCCTTCTCTCTTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC
 CTTCATGTGGCTCAGTGCCCTTCTCTCTTACCAGTCTATTTCCACTCTTTCACTGTGTCT
 GACATGTTGTGCTCTGTTCCATTTTAAACAAGTCTTACTTTCCAGTCTGTACAGAATG
 CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGGTGTGGCACTGGTGTCTGGAGACCTG
 GATTGTAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGAGCAAGGCATTTGGCTGCTGTA
 GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCACCTCACAGTGATG
 TTGTGGGGATCCAGTGAGATGAATACATGTAAGTGTGGTTTGTGAATTTAAAAAGTGCTAT
 ACTAAGGAAAGAATTGAGGAATTAACATACGTTTTTGGTGTGCTTTCAAATGTTTGA
 AAATAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFI LAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEGLWMSCVVSQSTGQI
QCKVFDSLNLNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDEDEVQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMTPVFNARYEFGQALFTGWAAASLCLLGALLCCSC
PRKTTSYPTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC
GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG
TGACCGCCCAGCCCATGTACGAGGGGCTGTGGATGTCNCGTGTCGCAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGC
GCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCATGACCCAGTCAATGCCAGGTACGAATTTGGTCA
GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCT
GTTCTGTCCC

FIGURE 100

ACCCTTGACCCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCCATCNTAGCACTGCCCTGCCCCAGTG
GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT
GGATGTCCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
TGATAGCAATCTNNNTGGCCACC GTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
GGTGCGAAGATGAGGATGGCTGTCTATTGGGGGCGCGATATTTCTTCTTGCGAGGTCTGGCTA
TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

FIGURE 101

GGGCCCACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTGCGAGANACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCCTTGGCATGAAGTGTATGAAGT
GCTTGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGGCGCGATATTCTT
CTTGCAAGTCTGGCTATTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTGCGA

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCAGGCCATTTACCGAGGGGCTTGGATGTCNT
GCNTGTGCGAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCCCTGTCCC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
21

FIGURE 103

AGCAATGCCCTGCCCCAGTGGAGGATTAATTCCTATGNTGGGACAAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTGCGAGAGCACCGGGCAGATCAGTGC
AAAGTNTTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCTCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTTGGAAGACGCTAGGTGTCAGAAGATGAGGATGGCTGTCAATGGGGGCGCATATTT
CTTNTTGCAGGTCTGCTCATTTAGTTGGCCACAGCATGGTATGGCAATAGATNGTTCAAGA
ATTTTATGACCCATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCTGCGCAACC

FIGURE 105

TCATAGGGGGGCGCGATATTTTTCTTGCAGGINTGGTTATTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTCTG

1056789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

FIGURE 106

TTCTTGGGATGGATCCGCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAAGTGATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTGTCATGGGGGCGCGATATTTCTTCTTGCGAGGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCACG

100
90
80
70
60
50
40
30
20
10
0

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGACGAGGGGCTGTGGATGTCNGCGTGTGCGAGAGCACCCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCCTTGGCATGAAGTGATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTGCGAA

FIGURE 108

GC GTG CCG CT CAG CT CG CCG GGC ACC GCG G C C T C G C C C T C G C C C C T C G C G C C T G C A C
C G C G T A G A C C G A C C C C C C C T C C A G C G C G C C C A C C C G G T A G A G G A C C C C C G C C C G
A C C G G T C C C C G C C T T T T T G T A A A A C T T A A A G C G G G C G C A G C A T T A A C G C T T C C C G C C C G G T
G A C T C T C A G G G G T C T C C C C G C C A A A G G T G C T C C G C G C T A A G G A A C **ATG** G G G A A G G T G G A G
C A G G T C C T G A G C C T C G A G C C G C A G C A C A G A G C T C A A A T T C C G A G G T C C C T T C A C C G A T G T T G T
C A C C A C A A C C T A A A G C T T G G C A A C C C G A C A G A C C G A A A T G T G T G T T T A A G G T G A A G A C T A
C A G C A C C A G T A G T A C T G T G T G A G G C C C A A C A G C G G A A T C A T C G A T G C A G G G G C C T C A A T T
A A T G T A T C T G T G A T G T T A C A G C C T T T C G A T T A T G A T C C C A A T G A G A A A G T A A C A C A A G T T
T A T G G T T C A G T C T A T G T T T G C T C C A A C T G A C A C T T C A G A T A T G G A A G C A G T A T G G A A G G A G G
C A A A A C C G G A A G A C C T T A T G G A T T C A A A A C T T A G A T G T G T G T T T G A A T T G C C A G C A G A G A A T
G A T A A C C A C A T G A T G T A G A A T A A A T A A A A T T A T A T C C A C A A C T G C A T C A A G A C A G A A A C
A C C A A T A G T G T C T A A G T C T C T G A G T T C T T C T T G G A T G A C A C C G A A G T T A A G A A G G T T A T G G
A A G A A T G T A A G A G G C T G C A A G G T G A A G T T C A G A G G C T A C G G A G G A G A A C A A G C A G T T C A A G
G A A G A A G A T G G A C T G C G G A T G A G G A A G A C A G T G C A G A G C A A C A G C C C A T T T C A G C A T T A G C
C C C A A C T G G G A A G G A A A G A G C C T T A G C A C C G G C T C T T G G C T C T G G T G G T T T G T T C T T T A
T C G T T G G T G T A A T T A T T G G G A A G A T T G C C T T **G T A G A G G T** A G C A T G C A C A G G A T G G T A A A T T G
G A T T G G T G G A T C C A C C A T A T C A T G G G A T T T A A A T T T A T C A T A A C C A T G T G T A A A A G A A A A T T
A A T G T A T G A T G A C A T C T C A C A G G T C T T G C C T T T A A A T T A C C C C T C C C T G C A C A C A C A T A C A C
A G A T A C A C A C A C A A A T A T A A T G T A A C G A T C T T T T A G A A A G T T A A A A T G T A T A G T A A C T G
A T T G A G G G G A A A A G A A T G A T C T T T A T T A A T G A C A A G G A A C C A A T G A T A A T G C C A C A A T
G G C A T A T T G T A A A T G T C A T T T T A A A C A T T G G T A G G C C T T G G T A C A T G A T G C T G G A T T A C C T C
T C T T A A A A T G A C A C C C T T C C T C G C C T G T T G G T G C T G G C C C T T G G G A G C T G G A G C C C A G C A T
G C T G G G A G T G C G G T C A G C T C C A C A C A G T A G T C C C C A C G T G G C C C A T C C C G C C C A G G C T G
C T T T C C G T G T C T T C A G T T C T G T C C A A G C C A T C A G C T C C T T G G G A C T G A T G A A C A G A G T C A G A
A G C C C A A A G G A A T T G C A C T G T G G C A G C A T C A G A C G T A C T C G T C A T A A G T G A G A G G C G T G T G T
T G A C T G A T T G A C C C A G C G C T T T G G A A T A A A T G G C A G T G C T T T G T T C A C T T A A A G G G A C C A A
G C T A A A T T T G A T T G G T T C A T G T A G T G A A G T C A A A C T G T T A T T C A G A G A T G T T T A A T G C A T A
T T T A C T T A T T T A A T G A T T T C A T C T C A T G T T T T C T T A T T G T C A C A A G A G T A C A G T T A A T G C
T G C G T G C T G C T G A A C T C T G T T G G G T G A A C T G G T A T T G C T G C T G G A G G G C T G T G G G C T C C T C T
G T C T C T G G A G A G T C T G G T C A T G T G G A G G T G G G G T T T A T T G G G A T G C T G G A A G A G A G C T G C C A
G G A A G T G T T T T T C T G G G T C A G T A A A T A A C A A C T G T C A T A G G G A G G G A A T T C T C A G T A G T G
A C A G T C A A C T C T A G G T T A C C T T T T T A A T G A A G A G T A G T C A G T C T T C T A G A T T G T T C T T A T A
C C A C C T C T C A C C A T T A C T C A C A C T T C C A G C G C C A G G T C C A A G T C T G A C C T G A C C T C C C C
T T G G G G A C C T A G C C T G G A G T C A G G A C A A A T G G A T C G G G C T C A G A G G G T T A G A A G C G A G G G C
A C C A G C A G T T G T G G G T G G G G A C A A G G G A A G A G A A A C T C T C A G C G A A T C C T T C T A G T A C
T A G T T G A G A G T T T G A C T G T A A T T A A T T T A T G C C A T A A A G A C C A A C C A G T T C T G T T T G A
C T A T G T A G C A T C T T G A A A G A A A A A T T A T A A T A A G C C C C A A A A T T A A G A A A

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFPTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMQLQPFDDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMSKLRGVFE
LPAENDKPHDVEINKIISTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLMRKTVQSNPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCACCTTTCAACCANTACTCACATTTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTTTGTCATTTTGTACATTGTCTATGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTAACTTTAAATGAGC

110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNNTTCCTCGCCTGTTGGTGCTGGCCNNTGGGGAGCTGGAGCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTGCTTT
CCGTGTCCTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAAGTCTTATTAGAGATGTTTAAATGCATATTTA
ACTTATTTTAAATGTATTTCATCTCATGTTTTCTTATTGTCAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGCTG

FIGURE 112

CCCTGGTGGTTTTGTCTTTAATTCTGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACNNGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNACACATACACAGATACACANANACAAATNTAATGTAAACGATNTTTTAG
AAAGTAAAAATGTATAGTAAC

Figure 1. The effect of the concentration of the *Agrobacterium* suspension on the transformation efficiency of *Agrobacterium* strains. The concentration of the *Agrobacterium* suspension was 10⁶ cells/ml (A), 10⁷ cells/ml (B), 10⁸ cells/ml (C), and 10⁹ cells/ml (D). The transformation efficiency was determined by the number of transformants per 10⁶ cells of the *Agrobacterium* suspension. The data are the mean \pm SD of three independent experiments.

GGTGGCCCATTCCTCCGCGCCAGGCTGCTTTCGGTNTTCAGTTCGTGCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCCAAGGAATTGCACGTGTGCAGCATNAGACGTAC
TTGNNATAAGTGAGAGGCGTGTGTGACTGATTGACCCAGCGCTTTTGAAATAAATGGCAGT
GCTTTGTTTCANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAAGT
TTATTCACAGAGATGTTTAATGCATATTTAAATTTAATGATATTTNATNTCATGTTTTCITA
TTGTCCACAGAGTGACGTTAATGCTGCGTCTGCTGAANTNTGTGGGTGAAGTGGTATGCG
TGCTGGAGGGCTGTGGCTCCTCTGTCTTTGGAGAGCTCTGGTCTAGTGGAGGTGGG

AAACCTTTTAAAGTTGAGGGGAAAAGAATGATCCTTTTATTAATGACAAGGGAACNCNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAATGACACCCCTTCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCGAGCATGTCTGGGGAGTGGGCTCTGCTCCACACAGTAGTCCCCANGTGGCCCANCCCC
GGCCAGGCTGCTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTGGAAATAAATGGCAGTGCTTTGTTCANTT
AAAGGGNCCAAGNTAAATTGTATTGGTTTCATGTAGTGAAGTCAAANTGTATTATTCAGAGATG
TTTAATGCATATTTAANTATTTAATGATATTTCACTCATGTGTTCTTATTGTCAAGGGT
ACAGTTAATGCTGCGTCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCAGGCTGCTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCCAAGGAATTGCACGTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGAAAATAAATGGC
AGTGCTTTGTTCACCTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAGAGATGTTTAAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTC
TTATTGTGACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNPFFSTAVKLSTGCSGILISQHVLTAAHCVHDGKDYVKGSKKLRV
GLLKMRNKS GGGKKRRGSKRSRREASGGDQREGTREHLQERAKGGRRRKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRPCSVSDESNDLLYQYCDAESGSTGSGVYLRLLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNAVVRITPLKYAICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCGTCCGGCAACTCTTGGCACACTGCTCCTCTTTCTGGCTTTT
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCTTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGAGGGGCCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAACTTTTGGATATGTGCATCAGTGGTTTATGCCAAATTTGTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCGCAACCAATCGGATGATACT
GTGGTTCAGCTTCCCTATGGAAGTAGACATATTCGCTCTGTCTTAAAGGTCTCGATCACTT
ATATCTGGAACCAAAACCCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCAGAGCAGATTTTCATTGTCAAGATTCTGTAATCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCTTCCCTCGTGGGAGGCCACCCCATGGACCGCGTGTCTCCTCC
TCGTGTGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGACATCCAGGGGCA
TGTCACTTCAGTGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCTCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAACTTCCAGTCGAGGCCAAGTGGCATGGTTCAAACAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTGACAGGAGCCCTCGTAAAGTTGTAAGAGCAGAGCTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGAATCATCTACCAAAGCTTTTTGGCTCTCAAATTAAGATTGATTAGTTTCAA
AAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLFLAFLLLSSRTARSEEDRDGLWDAGWPWSECSRTC GGGGASYSLRRCLS
SKSCEGRNIRYRTC SNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDFDNPCSLKCQ
AKGTTLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHRLVLKGPDHLYLETCTLQGTGKGENSLSTGTFL
VDNSSVDFKQFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFP
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDP CPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFKQAQEELEEGA AVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAATCCCGTGGAGGGGCGGTGGGCCCTCGGGCCTGAC
AGATGGCACTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGCGCTGTGGCTGGCG
GCCCGCCGGTTTCGTGGGGCCAGGGTCAGCGGCTGCGCAGAGCGGGGACCCCGGCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGCGAACAGCGGCTTGGGCCGCGCCACGGCCGCGG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCCGAGGAG
GCGGCGGGTCACTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGGCGCCAGAGCTTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGAGGG
ATCTTCCAGTGCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTACCAATCTTCTCCTTGGACTCCTCAAAGTTTCAGCTCCCAGCA
GGATTGTGGTAGTTTCTTCCAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAGCTTTTGTATAGCCGAGCAAAGTGGCTAACATTCTTTT
TACCAGGGAACTAGCCCGCGCTTAGAAGGCACAAATGTACCGTCAATGTGTGCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGTATGGGCTTTTTTCAAAGTCCAGTAGAAGGTGCCAGACTTCCATTTATTT
GGCCTCTTACCTGAGGTAGAAGGAGTGTGAGGAAGATACCTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAAGTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAATAGGGAACAAGGAGTAAAGAGCTGTTTATAAAAGTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAAGTGTACTTGAAGAAAAAGAAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATATTTT
TGGGATAAGAGAATTTGAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGTAAAATTATACTGGGCAAGCATGGATGACATATTA
ATATTGTGAGAATTAAAGTACTCAAAGTGCTATCGAGAGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTGTTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGAATAAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLGLLKSSAPSR
IVVVSCKLYKYGDINFDDLNSEQSYNKSFCYSRSLANILFTRELARRLEGTNVTVNVLHPG
IVRTNLGRHIHIPLLVKPLPNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

Figure 1 consists of seven sub-graphs, labeled (a) through (g), each showing the growth of a different bacterial strain as a function of temperature. The y-axis for all graphs is 'Growth' and the x-axis is 'Temperature (°C)'. The strains and their approximate peak growth temperatures are: (a) *E. coli* (37°C), (b) *S. aureus* (37°C), (c) *P. aeruginosa* (37°C), (d) *B. subtilis* (55°C), (e) *L. monocytogenes* (37°C), (f) *S. typhimurium* (37°C), and (g) *E. coli* O157:H7 (37°C). Each graph shows a characteristic bell-shaped curve where growth increases with temperature up to a point and then decreases as the temperature continues to rise.

GGGGATTGTAAGAGGAGAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAAAGGAACAAGGAGTAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGCGATTGAGAACTTGTTACT
TGAAGAAAAAGAAATTTTGATATTGGAATAGCCTGNTAAGAGGNCATGTGGGTATTTTGAG
TTACTGAAAAAATTATTTTGGGATAAGAGAATTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTTGTA AAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGA AATTAAAGTGACTCAAAGTGCTATCGAGAGTTT
TCTAAGTATCTTTGAGTTTATCGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTGTGGA AATTACTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGTGCCTGGAGAATCCTCCGCTGCCGTCCGGTCCCGGAGCCAGCC
CTTTCTTAACCCAAACCAACCTAGCCAGTCCCAGCCGCCAGCGCTGTCCCTGTACAGGAC
CCCAGCGTTACCATGTCATCCTGCCGTCTTCCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCTGGTAACTTGGGTTTTTACTCCTGTAACAACGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTGGGAAGCTTCCGATGTCAATTAAGGAAGAATT
TCCAAATGAAATCAAGTAGTGTGTCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCTCAAATTGTTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTGAGCGATCAGTGAAGCATTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTCAGAAATTGGGACTTAGCAGAAATCACCCTCTTGATCGCAGCAAAAGAA
ATATCATTTGGATATTTTGAAGAAAGGACTCGGACAACTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTCTGCAATTTGGGGATGTTTCAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCTCTTT
GTCCGAGAAATAACATTTGAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAGGTACAATAAACTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGTCTGTAAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAAATCTTGGAATACTCAAGCAATTCG
TATTTGACTTACATTCTGAAAACTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTAAAAAATTG
AAAAACAGTTTGAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVVFARVDCDQHS DIAQRYRISKYPTLKLFRNGMMMKEYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEEELTEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFHMYVFGDFKDVLI PGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
BQAQDVASSPPSSSQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCA GCGATCAGTGAAAGCATTGGCAGATTA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGCNGGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTGCGCTCCCGGAGCCAGCCCTTTCCTAACCC
AACCACACCTAGCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTNCGGANCCCAGCGTNACC
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACAACCTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTGTTTTAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAGCATTGGCAGATTACATCAGGC

bioRxiv preprint doi: <https://doi.org/10.1101/123456>; this version posted July 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 128

GCCCACGCGTCCGATGCGGTTACGTTTCGCGGCCTTCTGCTACATGCTGGCGTGCTGCTCA
CTGCGCGCTCATCTTCTTCGCCATTGGCACATTATAGCATTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAAACCCCTGAATCCCCCTGTACTCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGTCACTGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGTAGTGGCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTCCACATTTTGTGTTGTGAAAGACTG
TTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTGTGACAGGTTTGAACCTGCACTTCTTAAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCTAGTACATTGGAAGCTTTGTTTATAGGAACCTTGTA
GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTACAAAAATAAAAGCGGAATTTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATCTTAAGCATAAGTAAACATGATATAAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAGCTATTTTAAATGTGTTTTATTGTTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATCTTAAGAATTTGCAGG
TACTACAGATTTTCAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCTTTAGT
GCAATACAATAAACTCTGAAATTAAGACTC

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAAALIFFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPELLAYHIWRYMSRPVMSGPLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG
AATCCCCCTTGACTCCAGAGTACCTCATCCACGCTTCTTCTGTGCATGTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCCTCTTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCAGGACTCTATGACCCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTATCTTCTAGCATTTT
TTACTACCTATATGGCATGATCTATGTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 131

CGGACGCGTGGGGGAAACCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGCGCGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGAGGTTGGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAAGTTGACCTACCCCTTGACAC
CTACCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATGGAATGTGAATCTGCATGTACA
GAAGCATATTCCTAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAATAAGTTATATTCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCT
TGTCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAATTCACCTCCTCATAGAGCTTTTAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGS LWVRTQLGLPPLLLLTALAGSGTASAEAFDSVLGDTASCHRA CQLTYPLHTYP
KEEELYACQRCRLFSICQFVDDGIDLNRKLECESACTEAYSQSDEQYACHLG CQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFY LQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSY LQMRNSQAHRNFLEDGESDGF LRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVR SKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

TTGGGTGATACGGCGTCTTGCCACCGGGCTGTGAGTTGACCTACCCCTTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTGAGAGAGGTTGCAGGCTGTTTTCAATTGTGAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGTTGTCAGGAATCAGCTGCCATTTCG
TGAACTGAGACAAGAAACATTTATGTCCCTGATGCCAGGAATGCACCTACTCTTCTCCTAA
CTCTGGTGAGGTCATCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGCCCGGAGGTTCGGGGACCG
CTTCGGCTGAAGCATTTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGCCACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTGCTGAACTGAGACAAGAACAACCTTATGTCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLAGAVLLSSGHGEEQPPETAQAQRCFCQVSGYLDCTCDVETIDRFNNYRLF
PRLQKLLSDYFRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDSESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACATTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAAACAGCGATTTGATGGAATTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT
CTTNGAGCGCCCAGATTTTCAACTNNTTACTGGAAATAAAATTCAGGATGAGGNAAAACAAA
TGTTACTTTTGGAATACTTCATGAAATCAAGTCATTTCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATCTGGACAGAAAA
CGATCAGGCATGGAACCTCCCTTCGTCACTCACCTGTTCTTGCCCTGGTGTCTTGACAGG
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCTATTCCAGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCAATGCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTATCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGTATGGGGGATT
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCTAGAAAGCAAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCTGGGAGTAAGTATTTCCCCATCCCCAGGCCTGTGCCCTCTCTGGTCT
CGTGCTGTGGCAGCTCTGTCTTCAGTTCCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGAAGCCTGGCACCCTATGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAGCGGTATGATGCCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAAGCTCCATGTTTCTTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTGAGGAGTTCAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA
ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTCGCGTG
AGCCAAGATTGTGCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAA
AATAATAATAATAATAATTAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAATAAATGTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPFEEFGYSVLQHVGGGQRWMLVGAPW

DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTCCCCACCATCACTCNTCCCATTCCTTCCAAC TTATTTT TAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAAC TCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTCTTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCACAATGCCCCATGTGCCAAGGGCCACTAGGTGACTA
CCAACTGGGAAATTCATCTCATCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

FIGURE 141

AAAGTTACATTTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTGAGGCTTCGCTCGGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCCAGACTTTACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCATGTTTCTCAGATGAAGTGGCCATTCTGCCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTTTGATGTGGAGCCCAGTGATCGCGCTGGAGA
AACAGTGTACTATTCTGTGGAATACCGGGGGAGTACGAGAGCCTGTACAGAGCCACATCT
GGATCCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATAACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCAGA
CATTGCTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCTGGTACTGGCCCTGTTTGCCCTTGTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTGCTGTGAAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGGCCCGTGG
TGGTCTCCAGACACTTGAAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACCTCCTCAGGGCCTGGAT
CTC**ATAG**GTTTTCGGAAGGGCCAGGTGAAGCCGAGAACCTGGTGTGCATGACATGGAACCC
ATGAGGGGACAAGTTGTTTCTGTTTTTCCGCCACGGAACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACAGAGTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTGAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTCCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACCTTCAAGCCCAATGCCGTGCAGAGGGGAATGGCTTACGAGCTCTACAGT
AGGTGA CTTGGAGGAAGGTACAGCCCACTGAAAAATGGGATGTGCATGAACCGGAGGATC
CATGAACACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTAGCAAAATAAAAAGGGCCACCTTGGCCAAAAGCGGTAAAAAA
AAAAAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLFMWFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPISSWCSTLGEPECDVTDITATVPYNLRVRATLGSQTS
SAW
SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMILILV
VVPLFVWKMGRLQLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTGAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAAACCCCTTCGCGCGCAAGGG
GTTTNGCGAACCCTTCGCGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAGGAGAGAGGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACCGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTTTGGGAACTGGGACTCCCTGTGGGGAGGAGAGAAAGCTGGAAGTCCCT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGAGTGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGTTGGAGCTACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGTTGAATGCAGCGTGGAGTCTGTGTCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCATTAAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTAACAACCCGGCCGACATGTCTCCTTCTGCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCTCCTTTACAGCCACCGACTCAGTGAAGTGC GGCTGTCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACTTCAACCAGGAATCTACGGGAATTCAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTGTCAACGTTGCCAGTACCTCTAACCATTCTCT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCTACAAGAATGATGCCTACTTTC
TTCAAGACCTGAGCCTGGAGTCTCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCTGTCTCGAGACTGTACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCATCTCAGATCT
TCCAGAGCCTCAGCGTAACAGCCGGCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCAGGCCCACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTGCCTGAGACTCCCTTCGAGGATTGCACCGCCCGTCTCTAAGCCTC
CCCACAAGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLTYNTGRHVSFLPAPRPVVNVSGGFL
LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAAILSL
FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
TVTWILIDRALNITSLQMHSRLRLSQNPSPQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCGCCGAGCCTCGCGCGCGAGCCTC
 GTTCGTGTGTCGCCGCCCTCGCTCCCTGCGAGCTACTGCTCAGAAAAGCTGGGGGGGCCACCCCTGGCAGACTAACGAA
 CGAGCTCCCTCCCAACCCCACTCAGAGTCTAATTTTGAGCCTTTGCTCGCAATTCTTCAGGTTGAGGGGAG
 CGCAGAGCGGGAGGCTCGCGTATTCTTGCACTCAGCAACCCAGCTCGCCCCCGGAGCGCTCGGTGCTCAGGCCCTTC
 GCGAGCGGGGCTCTCGCTGCGGGTCCCTTTGGAAGGCTCTGGCGCGCTCAGAGGCGGCGCGCTCGCGTTGGCT
 CACCTCCCGAGGAACTTCACACTGGAGAGCCAAAGGAGTGGAGAGCCTCTCTTGAGAGATTTCCTGCGGGAA
 ATCCTGAGGTCATTCAATTGAGAGTGATCCGCGCGGGAGTGCTCAGAGTAACCAAGTCTGCTCATGCGCTAGA
 GCAATTCAGCCATGGTGGTTCCCAATGCCACTTTATTGGAGAACTTTGGAAAAATACATGGATGAGGATGCT
 GAGTGGTGGATGACCAACAACGAGGGGAAAGGGCCATCAGAGACAATGACATGCGAGAGTATTTGGACCTTCTAT
 AATAAATTACGAACTCAGGTGATTCCAAACAGCCTCTAATATGGAGTATGACATGGGATGTAGAGCTGGAAGA
 TCTGCAGAACTCTGGGCTGAAAGTGTCTTGTGGGAAACATGGACCTGCAAGCTTGCTCCATCAATTGGAAGAGAT
 TTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGTTTCATGTACAATCGTGGTATGATGAAGTGAAGACTTT
 AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTCAGGTGTTCTGGCCCTGTATGTACACATTATACA
 CAGGTCTGTGGGCACTAGTAAACAGAAATCGGTTGTGCCATTAATTTGTGTACAAATGAACTCTGGGGCAG
 ATATGGCCAAAGCTGTCTACTGGTGTCAATTACTCCCCAAAGGGAACCTGGTGGGGCCATGCCCTTACAAA
 CATGGGCGGCCCTGTTCTGCTTGCCCACTAGTTTGGAGGGGGCTGTAGAGAAATCTGTGCTACAAAGAGGG
 TCAGACAGGTATATCCCCCTGAGAAAGGAGAAACAAATGAAATGAGACGACAGCAGTCAAAAGTCCATGACACC
 CATGTCCGACAAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAATGTCCCAATTGTTTCT
 TGTGAAGTAAGATTAAGAGATCAGTGTCAAAGGAAACAACTGCAATAGGTACGAATGTCTGCTGGCTGTTGGAT
 ATAAAGCTAAAGTTATTGGCAGGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
 ATAAAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTCAATCAAGTCCCAATAGA
 AATGGTATTCAACAAATGGCAAAATCAGTCTGCTAATTCCTTCAAGTCTCAAAGTAACAGTTTCAGGCTGTG
 ACTTGTGAAACAGCTGTGAACACCTCTGTCCATTTCATAAGCCTGCTTCACATTGCCCAGAGATATCTGTCTC
 CGTAATCTGTATGCAAGCAATCCACATTATGCTGTGTAAATGGAACTCGAGTTTATTTCTGATCTGTCCAGTATC
 TGCAGAGCAAGCTACATGCTGGAGTGGTTCGAAATCAGCGTGGTTATGTTGATGTAACTGCTGTGACAAAGA
 AAGACCTACAGTCTCTCTTCTCAAGATGGAATCTTCTCAGAAAGTTTACAGAATCTCCAGGAGGAAGGCATT
 AGAGTGTGCTGTGTTGTTGGAAGTGAATCTTGGAAAGAGGACCATAAGAATATTCCAAATGCAATATTTCTGA
 ATTTTGTATAAACTGTAACTTACTGTACAGAGTACATCACTATTTTCAGCGCAATAGCTTTGGGAAAGTAAATGAAATATATGG
 TAAATCTTGATAAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAGTAAATGAAATATATATGG
 TTTTAGAAATCTGTGTTAAATATTGCTATATTTCTTAGCAGTATTTCTACAGTTAATTAATCATGATCATGAT
 GTTCTAAGTTTATATATATATATGTTGCTTTGTATATGCCAATAAATAAGTAAATCTAAACATTGAATGTGAATG
 GCCCTCAGAAATCATCTAGTGCAATTAATAAATAATCGACTTAAAACTGAAAGGAACCTTATCACATTTTCCCC
 AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAAATAATTTCCACTTAATGTAATGTAAGCTTTTTC
 TGTTAATTTAGGCATATAGAAATATTAATTTCTGATATTGCACTTCTATTTTATATAAAATAATCTTTAATATC
 CAAATGAATCTGTTAAAAAGTTTGTATCTCTTGGGAATGGCTTAAAAATAAATGTAATGAAGTCCAGAGTGGTGT
 ATGAAACATTCCTAGTGATCATGTAGTAATGTAGGGTTAAGCATGGACAGCCAGAGCTTCTATGTAATGTTA
 AAATTGAGGTACATATTTTCTTTGTATCTTGGCAATATCTCTGAGGCGGAGGAAGTATAATAGCAAAAGGT
 GAACAAAGTGAATTAATGTTATACATTAACATGCACTGATTTTAAAAATGTAATGACCTTGTATATA
 ATATTGCCATATCATGTGTACATAATGGTGATATATTTGTTTATGAAAAATGATATGCTGCTTGTATATAA
 AATCTGTAATGTTAGTTTGTGAATTTTTTCTGCTGGTGGATTTACATATTAATTTTTTCTGCTGGTGGGA
 TAAACATTAATAATCATGTTTCAAAAAAAGAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLPMARAIPAMVVPNATLLEKLLKYMDEDEGEWWIAKQRGKRAITDNDM
QSILDLHNKLR SQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPPASLLPSIGQNLGAHW
GRYRPPTTFHVQSWYDEVKDFSYPYEHECNFYCPFRCSGPVCTHYTQVVWATSNRIGCAINLC
HNMNIWGIWPKAVYLCNYSKGNWWGHAPYKHGRPCACPPSFGGGCRENL CYKEGSDRY
YPPREETNEIERQQSQVHDTHVTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNR
YEC PAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKH YFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCFPHKPASHCPRVYCP RNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSES LQNPPGGKAFRV
FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCGGGCGCCGCTCCGACGGGCCAGCGCCTCCCCATATGTCCCTGCTCCACGCGG
CGCCCTCGGTCAGCATGAGGCTCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAACCAAGTTGGGAGACTTGTGCAAGGA
CTTTGCAGATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTC
TTTCTCAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTCCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCAGAGCTGCTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCCTTTTAA
AAAAATGCTTTTTTGTATTGTGCCATACGTCACTATACATCTGAGCTTTTATAAGCGCCCGGA
GGAACAATGAGCTTGGTGACACATTTTCATTGCAGTGTTGCTCCATTCTAGCTTGGGAAGC
TTCGCGTTAGAGGTCTCGGCGCCTCGGCACAGCTGCCACGGGCTCTCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACCAAGTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCT
GGTTAATTTTACACACCCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGCTCTTCCCAACCTGAGGATTTCTGAAGGTTTACAGGTTCAATA
TTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACACAAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTTCCCATGTAAATCTTCAATGTTAAACAGTGCAGTCTCTTTTGAAGGCTAAGAT
GACCATGCGCCCTTCTCTGTACATATACCTTAAGAAGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGATTGTACTGCTGTATTATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTCTAAGATGGAAGTAATAAAATATATTGAAATGTAAAA
AAAAAAAAA

FIGURE 149

MSLLPRRAPFVSMRLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMWIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNANNEKRRVYEE

Signal sequence:

amino acids 1-34

MSLLPRRAPFVSMRLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMWIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNANNEKRRVYEE

FIGURE 151

MLNSVLLWLTLALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPIYASP
PTGERRFQPPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMPLIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASGYNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRI LADKVG CNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLG
V NQGEGLKFVDGI VDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHWAPAVAADLHAQYGSPTYFYAFYHHCCSEM KPSWADSAHGDEVPIV
FGIPMIGPTL FSCNFSKNDVMLS AVVMTYWTNFAKTGDPNQVPQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLEL VPHLHNLNEIFQYVSTTTKVP PPDMTS
FPYGTRRSPAKIWPTTKRPAITPANNPKH SKDPHKTGPEDTTVLIETKRDYSTE LSVTI AVG
ASLLFLNILAFAALY YKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLT CPPDYTLTLRRSPDDI PLMTPNTITMI PNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAGAATGCGCGGCGACTCTGGGACCCCTTGGGTCGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTTCGAGTACTTGAACCGGGAGCACTCGCTGTTCGAAGCC
CTACCAGGGTGTGGGCAAGGCAGTTCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCCTTACCCAGATATGCAAAGTAAACAGGGTGCCCTTGTGGAACCGG
GTGCCATGTTTCTTGAGAGACTGGGAGTTGCAGGTGCACCTTCAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCTCATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCTCGGTGATTCGCTACGTCGAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGAGTCCGCGCTGCC
CCGCGGCTACTACTTCCGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAAGTGCAGTGGAGAGAACCCAGAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCGGCC
CCTGAGTGGCCTGGCCCTCTTCTCATCGTCTTTTCTCCCTGGTGTTCCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCCTGCGCCACCCTTTTGTGACTGTCAACCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTGGGGACATATATTCTG
TCACTGGAGTTTGAATGCAGGACCCCGCATTCCCATGGTTGTGCATGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTTCCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAAC
AGAATTTCATAGCCCAGGCTGCCGTGTTGTTTGACTCAGAAGGCCCTTCTACTTCAGTTTTG
AATCCACAAAGAATTAATAACCTGGTAACACCACAGGCTTTCTGACCATCCATTGTTGGGTT
TTGCATTTGACCCAAACCTCTGCCTACCTGAGGAGCTTCTTTGGFAACCAGGATGGAACCT
TCTTCCCTGCCTTACCTTCCCTTCACTCCATTTCATTGTCCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATAGTGGCCCTAGGGAGATGGCTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGGAAGCATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAAGCTCCAGGTTTGATCAAACCAAAGCAACATTTGTCTAGTG
GTCTGACCATGTGGAGATGTTTCTGGAATGCTGAGACCTGCTAGCTGCATGTTTGTAGT
TAGCATTTTGGAAATCCACTTTGAGTGTGAAAGTGTAAAGGAAGCTTTCTTCAACACTT
GGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTCTTCTTAAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCTCGGAGAGGT
CACTGTCAATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCTTATTCCTGCTGCTTA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAATCAGTTACAGGCCAG
AGTCTCTCTGGAGGGCCTGGAAGTCTGAGTCTCTCATGAACTCTGTAGCCTAAATGAAAT
CTTAAATACCCGATGGAACCAAACCAAAAAAAGGGCGGCGCGACTAGAGTCTG
ACCTGCAGTAGGGATAACAGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDITYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFEFTVERTPEEEKLHRDVFLPSVDNMKLPENTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGCGCGCAGCGACGGAGCTGGGGCCGGCTGGGACCATTGGGCGTGAAGTCAACTACGGATCAGTCT
CTGATGGTGGGTCGTAACTCAGTGGGACCTCCAAGATTTCATGAAGAAAATCAGTTGTCTTCATTCAAGAAT
TGGGGTCGGCTCAGAATCTCGACCTGGTGAAGAAATCTGTTTCTAGAGAAGGTTAAATTAAGCCTGCAGTCT
GACATGTTCCCGATTGAGGTGAACCAATGAAGAGAAAATAGAATACTTAATAATGCTTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGGCTGCTGGCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACTGATCCCGGTGT
CGACTCCTAAGATGGAAATGAGTAGAAGAGTGAAGAGAGATCATGCCGACCTGTGTGACGAGAGCCCTCATGCC
CAGACCCCGTTTATGAAGCTCTTTTGTACTGCACATCCCGAGTGTGGCCGAGCGCAGCATGGAAGCTCATGCC
CGCATCATTTTAAGCTGGTCTCAGTGCATGTGTTTATTCCGACGAGAGACGGTACCCACTGTATGTCTATCCCA
AAACAAGCGACAGAAATAGCTGCATCTGTTGGCTTAACAGGAAAACCGTATCACCCAAAATGGAAGCTTTCA
TTAGTCACTGTCAAAGGATCCGAGCCTCTTTGAAAGCCCTTGAACCTTGTGCTCTTTACCCCAATCACC
CATTTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTGCAGAACCGGTGAGTGTGAGGGATA
TCTATCTAAAGAAAACAAAATCTCTGCCAATGATTGGTCTGCAGACAGCTCTATTATAGAGACCATGGGAAAA
GCCGGACCTCTACAAAGTGGGCTGGCTTGTCTTATGGCTTCTCCAGATTGTGATGGAAGAGATTATTTCA
GGCACCGCCAAGTGGCTGTTCTGCTCTGGAAGCTGCTATTGCCGGTAAGAAAACAGTATCTGGAAGAGGAGC
AGCGTCGTGATACCTCTACGTTTGAAGAACAGCCAGCTGGAGAGACCTACGGGGAGATGGCCAAGATCGTGG
ATGTCCCAACCAAGCAGCTTAGAGCTGCCAACCCATAGACTCCATGCTCTGCCACTCTGCCCAATGTGAGT
TTCCTGTACCAAGAAATGGCTGTGTGACATGGAGCATTCAAGGTAAATGAAGCCCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAAGAAATGTACTTGGGTATTTCTCTCTGGGTGCCACCCCATCTGAACCAAAACCATCG
GCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAGAGCTCTTGGCCCTCTACTCTGCTCATGATGTCACTCTGT
CACCAGTTCTCAGTGCCTTGGGCTTTTGAAGCCAGGTTCCTCAAGGTTGCAGCCAGGTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCAGCTGAACATTCCGTCCGGATTCTTACAATGGCGTCGATGTACATTTCCACACT
CTTTCTGCCAAGAACCCACACAAGCGTTCTCCCAAGCCATGTGCCCGCTTGAAAATCTGGTCCGCTTTGTGAAA
GGGACATGTTTGTAGCCTGGGTGGCAGTGGTACAAATATTATGATGCATGTACAGGGAAGGATTCTAAAGG
TATGCAGTACAGCATATAGAATCCATGCCAATACAGAGCATAGGGAAGGTCCACTTCTAGTTTGTCTGTATAC
TAGGGCTAGAAGATTATGTCTTTTAAAGGCTAAATATTGTTTGGGGAACCCAGATGGTGGGGTTGAACACT
AAGCACTGTGCTGCAATGGTACGTGAATTGCTTGTGACAAATGGCCAGTTCAACAGGAATACAGAGGTACTT
TATCATGGCCAGCATCTGCTTAGAATGCCAGAAATATATAGTTCAAGACCTGGAAGTTGCCAATCCAAAGTTGCAC
TCTCTGCGCTGCCCATGTTACTATGTGATGGAACCCAGCACCTACACAAATTTTAAATCTTTAGACATT
TTTACCTTGTCTTGTAAAGATTCTTGAAGTGATTATCTAAATTAAGGTTGGCAACTTTTCTCTGAAGG
GCCAGTTTAAATATTTCAGACTGTGTGACCAAAAGCCACATACAGTCTCTGTCTAACTACTCACTCAATGTG
TTCGAAGCAGGAAGGCCACACAGACAGTACATAAAGCAATATGTGTAGCTGGTTCAGGCCAGACACAAACA
GATGGTGACAGCACTTGGCCCTGGGCTGTAGTTTGTCTGACCCCTCATCTAAAAATAGCCTATACATCAATTCG
CATCTCAGCACTTTGAGAACGAGTGAATACCAAGAAATATTCAATGGTCTCTCAGTAACTTCTGTAGAACAC
CAGAATTTGCTGTATCTGACACTAGAACAAAACCTGAGGGTAAATAAACATTAATAGAAATGAAATATTGTT
AACTGATTAGAAGAATACCTTGATGTTTATGATGATTGTGGTACAGATAGTTTAAAGTATGTTCTAAATATTGTT
CTGCTGTAGTCTATTGTCTGATATGTGAAATTTTGTATGCCATTATGATTTTATAGTTTAGGAAAAATTT
TTCTAAGACAGTTTATGATGACTCTATTCTGTAGTAATATTCAATTGCTGTACCTGCTTGGTGGTGAAG
GAGGCTAGAAGATGAATTCAGGCATTTCTTCCAATAAAACATAATATGGCTCACTCCCTTGGACAAGCTGTAGA
ACTGGATTCAATTTTAAACCATTTTATCAGTTTCAATGGTAAATTTCTGATTGATTTTAAATGCGTTTGTGGA
TGAATCTTGTCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTATATATTAGAAGCAATATAATATACATCTC
GTATTTCTCACTAATGGTGTCTAATTCAGAGAAATGAAAGTGAAGTGAGATCTCTGTGTCTATCGGCATTTCC
AACTTTTCTCTTTGTTTGTGTCAGTGTGCATTGAATATGTCTGTTTCTATAATAAATTTTTTAAGAATAA

><subunit 1 of 1, 480 aa, 1 stop

MLFRNRFLLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTPDPVY
EALLYCNI PSVAERSMEGHAPHHFKLVSVHVFIRHGDYRPLYVI PKTKRPEIDCTLVANKRP
YHPKLEAFISHMSKSGSGASFSPINSLPLYPNHPLCEMGEILTQTGVVQHLQNGQLLRDYLK
KHKLLPNDWADQLYLETTGKSRTLQSGGLALLYGFLPDPFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVP TKQLRAANPIDSM LCHFC HNVSFPC
TRNGCVDEMEHFVKIKTHQIEDERERREKKLYFGYSL LGAHPI NLTIGRMQRATEGRKEELF
ALYSAHDVTLSPLVSLALGSEARFPRFAARLIFELWQDHPKSEHSVRILYNGVDVTFHTSF
CODHHKRSFKPMCPLENLRFVFKRDMFVALGGSGTNYDADACHREGF

Signal sequence:

amino acids 1-18

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGLLWETGCTQIRYSVP EELEKGSRVGDISRD LGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDIIMEDKVKIYGVVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELV LKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLVNADTNDNPPVFPQASYSAYIPENNPRGVSLSVTAHDPDCENAIQITYSLAENTI
QGASLSSYVSINSDTGVLVALSSFDYEQFRDLQVKVMARDNGHPPLSSNVLSLFLVLDQNDN
APEILYPALPTDGGTGVELAPRSAEPGYLVTKVAVDRDRSGQNAWL SYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTITVAVADSIPQVLADLGSLESPA
NSETSDLTLYLVVAVAVSCVFLAFVILLALLRLRRWHKSRLLQASGGGLTGAPASHFVGVD
GVQAF LQTY SHEVSLTTDSRKSHLIFPQPNYADMLVSVQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCSVSVQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGS LTPVIPVLWEAEAGGSPEVGS LRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCGCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGCTCTTGCTGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGTCATGAGTGCCAACCCATTGCGAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTT
ACAGCTGCCCCTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCAACCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCGAGAGAAATTTTCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACCTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGGQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDFCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCCG
CGCGCCGCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCGGGCCCGCGCCCCG
GCCCGCGCCAGGTGAGCGCTCCGCCCGCGCGAGGCCCGCCCGGCCCGCCCCCGCCCCG
CCCCGGCCGGCGGGGAACCGGGCGGATTCTCGCGCGTCAAACCACTGATCCCATAAAAAC
ATTATCTCTCCCGGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCGCGCCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCCCGGGCCGTAGCGCGCGCGCCTGGA
TGCGGACCCGGCCGCGGGGAGACGGGCGCCCCGCCGAAACGACTTTCAGTCCCCGACGCGC
CCCCCCAAACCTTACGATGAAGAGGGGCTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGCA CGGCAACCGCATCTCGCATGTGCCAGTGCACGCTTC
CGTGCTGCGCAACCTCACCATCTGTGGCTGCACTCGAATGTGCTGGCCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCTGCCACATTCACGGCCTGGGCCGCCCTACACACGCTGCACCTGGACCGC
TGCGGCTGCAGGAGCTGGGCCCGGGGCTGTTCGCGGCCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTTCTCTGCACGGCAACCGCATCTCCAGCGTGCCGAGCGCGCCTTCCGTGGGCTG
CACAGCTCGACCGTCTCTACTGCACAGAACCGCGTGGCCCATGTGCACCCGATGCCCTT
CCGTGACCTTGGCCGCCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCACTCTGGGCTGGCTGCAGAAAGTTCGCGGCTCCTCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCTGGCTGGCCGCTGACCTCAAACGCCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCAC
GATGAGGAGCCGTGGGGCTTCCAAAGTGTGCCAGCCAGTGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCACGGCACATCAATGACTCACCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCGGGTT
CCCCACTCTGGGCCCTCGCGGAGGCCAGGCTGTTCACGCAAGAACCGCACCCGACGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCTGTGTAACCCCGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGTCTCTCTCCACGCGGCCAAGCCAGCCGGCGGCCGACCCGTGGGCGAGGCCAGGCCAG
GTCCTCCTGATGGACGCGCTGCCGCCGCCACCCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTGTTCAGAACCGCGCCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAAA
AAAA

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRLLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFACRNLTIILWLSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCGLQELGPGLFRLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRLHSLDRLLHQNRAHVHHPHAFRDLGRMLTYLFANNLSALPTEALAP
LRALQYLRINDNPWVCDRCRARPLWAWLQKFRGSSSEVPFCSLPQRLAGRDLKRLAANDLQGCA
VATGPYPHIWTGRATDEEPLGLPKCCQPDAAKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSESGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

GGAAGTCCACGGGGAGCTTGGATGCCAAGGGGAGGACGGCTGGGTCTCTGGAGAGGACTAC
 TCATCTGGCATATTTCTGAGGATATCTGTAGTAATACCAACAGCGCTCAGATATCGGGGACTTAC
 AGTCCACAGAAACCGCTCTCCGAGGAGCTGTAATCTCAGCAAGAACATGAGGCGACGGCGAA
 AGCTCTGTTGCAGACAAAGGCAGTCTCTTTTCTCTCTTTTGGGCTGATCTCTGGCG
 GCGCGCGGCGAAGCTAGAGCTATTTCTGTGGTGGAGAAACTGAGGCGAGCTCTTTGTGCAC
 CAATTTAGCAAAAGGACTCTGGGTCTGGAGCAGAGGGAATTTCTCAGGCGGGGGTTAGGTTGT
 TTTCCAGAGGGAAACAACATCACTTTGCAGCTCAATCAGGAGACCGCGGATTTGTTGTCTAAAT
 GAGAAATTGGACCGTGAGGATCTGTGCGGTGCACACAGACCGCTGTGTCTAGCTTTTCCAAG
 GTTGCTAGAGAGTCCCTTCGAGTTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGAC
 ACTCTCTCAGTATTTCTTGCAACAAACAGATGTTGGTGAAGTAGTACAGAGAGCAGTCTCTCGGG
 ACTAGCTTTTCTCTGAGGAATGCGGAAGACTTAGTAGTGAGGCCAAACAAATATTTGAGAACTA
 TATAATCAGCCCCAATCTCTATTTCTGGGTCTCTACCCGCAACAGCAGTGATGGCAGGAAT
 ACCCAGAGCTGGTCTGGACAAAGCGCTGGACCGGAGGAAGAGCTGAGCTCAGGTTAAACA
 CTCACAGCATGTGATGGTGGCTCTCGGCCAGATCTGGCAGTGCTCAGGTTACATCGAAGT
 CTCTGGATGTCAACGATATGCCCTGGAATTTGAGCAGCTCTTCTATAGAGTGTGCAGATCTCTG
 AGGACAGTCCGGTAGGCTCTCTGGTTGTGAAGTCTCTGCCCAGGATGTAGACACAGGAGTC
 AACGGAGAGATTTCTATTCACTTTTTCCAAGCTTCAGAAGAGATTGGCAAACCTTTAAGAT
 CAATCCCTTTGACAGGAGAAATGTAATCAAAAACCAATCGATTTGCAAAAACCTTCAGTCTCT
 ATGAAGTCAATATTGAGGCAAGAGATCTGGAACCTTTTCTGAAAATGACCGGTTCTGATT
 CAAGTGATAGATGTGAACACCATCGCCCAAGAGTTACCATGTCTCGATTATCCAGCCCAAT
 ACCTGAGAGACCGCTGAAACTGTGGTTGCACTTTTCACTGTTTTCAGATCTTGATTTCAGAG
 AAAATGGGAAAAATTAGTTGCTCATTTCAAGAGGATTTACCTTCTCTGAAATCCGCGGAA
 AACTTTTACACCTCACTAACCGGAGAGAACCTAGACAGAGAAAGCAGAGCGGAATACAACAT
 CATATCATCTGTCACTGACTTTGGGAGACCCATATGCTGATAAACAGCATCAATATAGACCGTGC
 TGATCGCCGATGTCAATGACAACGCTCCCGCTTCCACCCAAACCTCTACACCTTCTGCTCT
 CGCGAGACAACACAGCCCGCCCTGCACATCCGCGAGCGTACAGCGCTACAGACAGAGACTCAGG
 CACCAACGCGCCAGGTCACTACTCGTGTGTCGCGCCGAGGACCGGACCTGCGCCCTCACAT
 CCTTGGTCTTCATCAACCGGGACAACGGCCACCTGTTCGCCCTCAGGTTCTCTGACTACGAG
 GCGCTTGCAGGGTTTCCAGTTCCGCTGCGGTGGCGCTTCAGACAGGCTCCCGGCGCTGAGCAG
 CGAGGCGCTGGTGTGCGCTGGTGGTGTGAGCGCCAAACGACAACCTCGCCCTTCGTGCTGTACC
 CGCTGCAGAAACGCTCCGCGCCCTGCACGAGCTGGTGTCGCGCGGCGCGGCGAGCGGGGTAC
 CTGGTGACAAGAGTGGTGGCGGTGAGCGGCACTCGGCGCAGAACGCTGCTGCTGCTACCA
 GCTGCTCAAGGCCACAGGAGCTCGGTCTGTTGCGGCTGTGGGCGCAACATGCGCAGGTTGCGCA
 CGCCGAGGCTCTGAGCGAGCGCGAGCGGCGCAAGCAGAGGCTGGTGGTGTGTTCAAGAC
 AATGCGAGCGCTCCGCGCTCGGCAACCGCCACGCTGCAAGTGCTCTTGGTGGACGGCTTCTC
 CCAGCCCTTACCTGCTCTTCCGGCGCGCGGCGCGGACCGAGCCGAGCCGACCTGCTGCTCCGCT
 TCTACCTGTGGTGGGCGCTTGGCTCGGTGTTCTGCTCTTCTCTTTTCGTTGCTGCTCTGTGTT
 GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCTCGGTGGGTGCTGCTTGGTGGCCGA
 GGGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCGACCGGACCCCTACCGAGAGCTACC
 AGTATGAGGTTGTCTGCGAGAGGCTCAGGGAACCAATGAGTTCAAGTTCTGAGCGGAT
 ATCCCCAATCTCCCTCCCAATGCTCCCTGGGAAAGAAATACAGGAATTTCACTTCCCCAA
 TAACTTTGGTTCATATTTCACTGAGCATATGACTTTTACATTCATAGATTTTATTT
 GTGGCATTTTCCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
 TTACTCTTGATTTTTCTCATGTTCTTTCTCCCTTTGTTTAAAGTGAACTTTACCTTTAT
 CCTGTCTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRGKHLHLQLNQETADLLNEKLDREDLCGHTEPCVLRQVLLSPFEFFQAEQV
IDINDHSPVFLDKQMLVKVSESSPPGTTTFLPKNAEDLDVGQNNIENYIIISPNYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLTTERPLDRESAEYNTITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFQTSTYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPPLTSLVSINADNGHLFALRSLDYEALQGFQFRVVGASDHGSPALSSEALVRVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGFLGVWAH
NGEVRTARLLSERDAAKHRLVLVLKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFSLFVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site,

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCCCACGCGTCCGCCCCACGCGTCCGCGCGTAGCCGTGC
 GCCGATTGCTCTCGGCCCTGGGCAATGTTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
 GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGCGTGG
 AGGTTGACAGAGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCCTGCTCACCTCTCCAGGTG
 GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
 AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCATCACATGGTGATGCTGTCTG
 TGATTCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGCGCTCACCTGTGGTGCT
 GGAGGAGCGGAGGACTCAAGGTGCAACGTCGAGAGAGCCTTTTCTCTGATGGCGCTGG
 AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
 CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACCTGTGAG
 GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTACAGGACCT
 TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCTGTTTACACCCCTGGT
 GCCGCTTTTCTGCAGTTGGCCCCCTCACTTTAACTCTCTGCCCGGGCATTTCCAGCTCTT
 CACTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACAGGTTTGGCACCGTAGC
 GTTCTCTAATATTTTATTATTTCAGGAGCTAAACCAATGGCCAGATTTAATCATACAGATC
 GAACACTGGAAACACTGAAAACTTCATTTTTTAATCAGACAGGTATAGAAGCCAAGAAGAA
 GTGGTGGTAACTCAAGCCGACCAATAGGCCCTCTCCAGCACTTTGATAAAAAGTGTGGA
 CTGGTTGCTTGATTTTCCTTATTCCTTTTAAATTAGTTTTATTATGTATGCTACCATTCGAA
 CTGAGAGTATTCGGTGGCTAATTCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
 GAAAGAAGTTGGAAGAGGAACCTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
 CATTTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAGAAGATCATTTGTTGAA
 CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
 AAAAATATTCAATAG

1051710501

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRPRPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETCLKIFIFNQGTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFLLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCCGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCTCGCCTCGCACTGCTGCGCTGGGCTCGGCGCGGGCTCTTCCTC
TTTGCCAGCCCCGACTTCTCTACAGCGCAGCAATTGCAAGCCCATCCCGGTCAAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCTGTGCTCGCTCTTCGCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCTGGCCCGACATGCTTGAGTGCAGCGTTTCCCCCAGGACAACGACCTT
TGCAATCCCCCTCGCTAGCAGCGACCCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAATGATGATGACAACGACATAATGGAACGCTTTGTAAAAATG
ATTTTGCCTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTA
GTCCCGGCATCTCTGATGGCTCCGACAGGCCTGTCCAGAGCACGGCTGACCATTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCAAGCACACTCTAGCTGTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCACCCGAATCTGTAGAAATATTCAAACATAATA
AAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETIQCHSLCVQVKDR
CAPVMSAFGFPPDMLCEDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIM
ETLCKNDFALKIKIVKEITYINRDTKIILETKSKTIYKLVGVSEKDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQGGELVITSVKRWQKQREFKRISRSIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACGATGGCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGACCTACGCGCGCCGCTGGGTGTTCTGTCTCGCGATC
AGCCTGTCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGGTGACGTCA
TGCTGAGGACTTGGTCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGGC
ACCATCCTGGGTGCGTGGCTGAACCTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTGCAACCCTCTGGGCGTCTCTGTGGCCAATGTGCT
GTCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCGGTAAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTGTCTGCCTGTGTGCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACCCCG
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTATCTGGCTGTGTGCTTTGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCAGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCTTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTTGCCCTGGTGTCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGECTTCTCGGTGGGCCCGGTGCCATGGAGTTGGCGGTGCA
GTGTTCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGTCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCGGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTGGGGCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCGCCCTCCGCGCCCGG
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTTGTAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRITYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFVGAIIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGVLVANVLSPVLV
KKGEDIPLMLGVYTI PAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCFLSLACVPFALVSQQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGILIMLMTALT VRRSEPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPPTSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSFPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 170

GTCCACATCTCGTCAACTGGGTCAAGTCCCTCTTAGACCAGCTCTTGTCCATTCATTGTGTAAGTGGACCAAC
 TAGTTCCTCCAGTAGGGGTCTCCCTCGGCAATTTCTTGATCGGCGTTGGACATCTCAGATCGCTTCCAATGAAGA
 TGGCTTGCTGCTTGGGTCTCGTGTGTTTCATAATCATCTAACTATGGGACAGAGTTGTGCGGGCAGCTCTGGGG
 AAGGACACCGGGCTGATCAAGCATCCAGGAAACACTCGAGGAGCTTGTCCAGCCTTGAAGAAGCTCTAGTGGTT
 TCTGAATCTAGCCCACTGGCGTAACTAGATGCAACTTCTGCAACTTCTGCTGGGGCTTGGGGCCAGGTGG
 CTACTTATTTCTTTTAGGGGATGTGTCAGGAGGTGACCACTCTCAGCGTGAATAACAAGTGTCCAGGAGAGTGC
 ATCTGTCAGTGTATCGGGAAGCTGTCCAGGAACCTGGGCGGGAGGAGAGCGGAGCAAGCTGGGGCCGCTTT
 CCAGGTGTTGTAGCTGTCTCAGGCGTCCCAATTCAGTGGACCTCTGAGGAAGGCTTGCTCAGCACAGGCGGG
 GCTGGATCGAGAGCAGTCTGCCACAGTGGGATCCCTGCTGTTCTTCTTGATGTGCTTGGCCACAGGGATTT
 GGCTCTTGATCCATGTGAGAGTCCAAGTGTGGACATCAATGACCACAGCAGCTTTCCTCAAGAGCGAGGAGA
 GCTGGAAATCTCTGAGAGCGCTCTCTGCGAACCCGGATCCCCCTGGACAGAGCTCTTGAGCCAGACAGGGCCC
 TAAACCCCTGCACACCTACAATCTGTCTCCAGTGAGCACTTTGGCTTGGATGTCAATGTGGGGCCCTGATGAGAC
 CAAACATGCAGAACTCATAGTGGTAGGAGCTGGACAGGGAATCCATTCTTTTGATCTGGTGTAACTGCT
 CTATGCAATGGGAACCCCCCAAGTCAGGTACCGACTTGTCAAGGTCAACGCTTTGGACTCCAATGACAAATG
 CCGTGGTGTGCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAATCT
 GACCGCCACAGACCTTGACCAAGGCCCAATGGGGAGTGGAGTTCTTCTCAGTAGACACATGCCCTCAGAGGT
 GTGGACACCTTCAGTATTGATGCAAGACAGGCGAGGTCAATCTGCGTGCAGCTCTAGACATGAAAGAAACCC
 TGCTACGAGGTGGATGTTCAGGCAAGGGACCTGGGTCCCAATCCCTATCCGAGCCCAITGGCAAAGTTCTCATCAA
 GGTCTTGATGTCTCAATGACAACTCCCAAGCATCCAGCTCACATGGGCGCTCCAGCCATCACTGGTGTGAGAAGC
 TCTTCCCAAGGACAGTTTATTATGCTCTTGTCATGGCAGATGACTTGGATTACAGGACAAATGGTTTGGTCCAATG
 CTGGCTGAGCCAAAGAGCTGGGCCACTTCAGGCTGAAAGAACTAATGGCAACACATACATTTGCTAACCAATGC
 CACACTTGACAGAGAGCAGTGGCCCAATATACCCCTCACTCTGTAGCCCAAGACCAAGGACTCAGCCCTTATC
 AGCCAAAGAAACAGCTCAGACTTCAGATCAGTGACATCAACGCAATGCGACCTGTGTTGAGAAAAGCAGGATAGA
 AGTCTCCACGGCGGGAACAACTTACCTCTCTTCACTCAATTACCAAGGCTCATGATGCGAGACTTGGGCAT
 TAATGGAAGTCTCTATACCCGCTACAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
 GGTCAAGTCTCAGAGTCTGACTGAATCTAAGAGATGGCGCGCTTTGAGTTCAGGTGTGCGAGAGGACGCGG
 GCAACCCATGCTTGCACTCAGTGTCTCTGTGGGTGACGCTCTTGGATGCCAATGATATAATGCCCCAGAGGTGGT
 CACGCTGTGCTCAGCGATGGAAGCAAGCCGCTCTCCGTGCTTGTGATGCTCATAGTGCCCTCAGTGGTGCC
 CACTGAGACTCCCAATGGCTTGGGCCACGCGGCACTGACACACCTCACTGGCCACTCAGAGCTCCCGGCCATT
 CCTTTTGACCAACATTGTGGCAAGAGATGCAGACTCGGGGCAATGGAGAGCCCTCTACAGACTCCCAATGG
 AAATGAAGCCCACTTCTATCTCAACCTCATACGGGGCAGCTGTTCGTCAATGTCCCAATGTCCGACGACT
 CATTGGAGAGTAGTGGAGCTGGAGATAGCGATAGTAGAGACCAAGGAAGCCCCCTTACAGACCCGAGCCCTGTT
 GAGGGTCACTGTTTGTCAACAGTGTGACCACTGAGGGACTCAGCCCGCAAGCTGGGGCGTTCGAGCAATGCTGAT
 GCTGACGCTGATCTGCTGCTGTACTCTTGGGCATCTTCGGGTGATCTGGCTTGTTCATGTCCTATCGCTGCG
 GACAGAAAGGAGGACACAGGGCTTACAACCTGTGCGGAGCGCCAGTCCACTACCGCCAGCAGCCCAAGAGGCC
 CCAGAAACACATTACAGAAGTGAGACATCCACTCTGCTGCTGTCTCAGGGTCAAGCAGCTGAGCCTTGTAAGT
 CGGGCAGTCCCAAGAGATGTGGACAAGGGGCGATGATGGAAGCAGGCTGGAGCCCTGCTGCAAGCCCTCT
 CCACCTCACCCCGACCTGTACAGAGCAGCTGCGTAAATCAAGGCAACAGGGGACAGCCCGCGAGAGCCGAGAGGT
 GCTGCAAGACAGCGTCAACTCTTTTCAACCATCCAGGCGAGGAGTGCCTCCCGGGAAGACTTGAACCTTCC
 CGAGCCCCAGCTTGCCACAGGCCAGCCACGTTCCAGGCTCTGAAGGTTTCAGGCGAGCCCAAGAGGAGGCTGCG
 TGGAGACAGGGCAGTGGAGGAGGCCCAAGAGGCCACAGGCTCTCTCTCAGACCTTGAGACGGCAGCACTCAT
 CAATGCGCAAGTGTCCCTTGAGAAAGATCAGGGCCCCGTGAGCTCTCGGGAGCGTGTGCGGCTGTCTGTGGT
 TGCTTTCGCGAGCGGGAACCCGTGGAGGAGCTCACTGTGGATTCTCTCTGTTCAGCAAACTTCCAGCTGTCT
 GTCCTTGCTGCACTCAGGGCCAATTCCAGGCCAAACCAACCAACGAGGAAATAGTACTTGGCCAGCCAGGAGG
 CAGCAGGATGAGGATCCAGGACAGAGTGGCCCAAGTGCAGGGCTGGAGGCCAGAGCAGACGCAAGAGGAGA
 AGGGCGTTTGGATCTGAAAGGAGCCTCTCTGTGAAGCACTGCTAGAAGAGAGAGCTGAAGTCTGCTGTGGAC
 CAGCACAGTCTGCGGCTTGACCGGCTGAGCGCCCTGACCGGCTGGATGCGAGACTCTTCTTGGCCCTCAC
 CACCAACTACCTGACAAATGTGATCTCCCCGGATGCTGACGACAGGAGGAGCGAGGACCTTCAGAGCGTTTCG
 CAAGGACAGGSCACCAAGAGCTGAGGCCCAAGGCAAGGCAAGGCTGGCCAGCACTTGTCTCGGAGATGACTCAT
 CTTGAGGATGCTGCTGGAACAGCGCTCCAGCATGCGGTTGGAGGCGCCTCCGAGGCGCTGCGCGGCTCTCGGT
 CTGCGGAGGAGCCCTCAAGTTAGACTTGGCCACAGTGCAGCCTCAGGCATGGAAGTGCAGGGGACCAAGTGG
 AAAGACGGGGCTGAGGGCAAGGACAGGCGAGCAGCAGCAGCAGGAGTGTGGAACATCTCCAGACGCTCT
 CTGATCTCAAGAACCCAGGGGCTGAGGATCTGTGGACAAGAGCTTGTTCATAATCTTGTAATCTCACTAGCTAG
 CGGCGGCTTGAGAACTTTAGGCTGACTGATGCTACCCCCACAGGAGGAGCAAGGCCAGGACTCAACAGCTGAC
 TGCCAAAGCAGCCCTTGTGAGCAGCTCTGAGTCTTTTGGAGGACAGGACGTTTGTGGCTGAGATAGTGTGT
 TCCCTGGCAAAACATATGTGGAGCAGAAAGGTGAGTCTCTGGCAGAACATGCCACGAGTATACAGGCGAGG
 AAAGGGTGGCCTTCTGGGTAGCAGAGTCAAGGGCTGTACCTCTGGGGTGGCAGGAATGCTCTCTGACCTAT
 CAATAAGGAAAGCAGTAAAAAAGAAAAAAAAAAAAA

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLGLLGGPGGYFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEGLLSTGRRLDREQLCRQWDPCLVSFVLDLATGDLALIHVEIQ
VLDINDHQPFRPKGEQELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVI
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSDNNSPAPFAESS
LALAIQEDAAPGTLTIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGNPIPAHCKVLIKVLVDNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDRQWPKYTLTLTLLAQD
QGLQLPSAKQLSLIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIIKAHADDLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMCLASSVSVVWSLLDA
NDNAPEVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTTPPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHQKADIHLVPVLRGQAGEPCEVGQSHKVDVKEAMMEA
GWDPCQLQAPFHLTPTLTYRTLNRQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLEP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSVLRLSVAAFAERNPVEELTVDSFPVQQISQLLSLLHQGFQPKPNHRGNKYLAKPGGS
RSAIPDPTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEBELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNRYNDNVISPDAAATEEPRTFTQTFGKAEPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAAAGMKVQGDPGGKTGTGEGSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 173

MSFLIDSSIMITSQILFFFGWLFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLILVFMVPPYIGYFIVSNIRLLHKQRLLFSCLLWLTFFMYFF
WKLGDPPILSPKHGILSIEQLISRVGVIQVTLALLSGFGAVNCPYTYMSYFLRNVTDDI
LALERLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDALEELSRQLFLETADLYATKERIEYSKTFKGYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTPDPTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAIS
SKSSNVIVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCGANACTATTTTTTGGATTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATT
CTTGCAACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTC
TTTNTTGAATTCGGGTTNNGNATACCTTCCAGAAAATATTTTTGGATTTGGGGTAGNTT
TTTTTCATGCGCCAATGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCCTGCTGATC
CTGGTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCATTC

175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLEVLLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGFLACPLSLEETDCYRVDIDQADMQKESKENQWL
GVSVRSQGGPKIVTCAHRYEARQVRVDQILETRDMIGRCFVLSQDLAIRDDELGGGWKFCFG
RPQGEHQFGFCQQGTAFAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDDGPPYEA
GGEKEQDPRLLIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSFGFGYSLAVADLNSDGWPDLLVGPYFFERQEEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAVGAPFDGDKGVFIYHGSSLGVAKPQSVLE
GEAVGIKSPGYSLSGSLDMDGNQYPDLLVGSGLADTAVLFRARPI LHVSHVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFYIAVPSSYSPTVALDYVLDAITDRRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGGGLPPVAP
ILNAHQPSQRAEIHFLKQCGGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGGPVIGLELMVTNLPSPDPAQFPADGDDAHEAQLLVMLPDSLHYSQVGRALDPAEKPLCLSN
ENASHVECELGNPMKRGAVQVTFYLLISTSGISIEETLEVELELLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGSRLRTLSGSAFLNIM
WPHEIANGKWLLYPMQVELEGGQGGQKGLCSRPFNILHLDVDSRDRRRRLEPEPQEPGE
RQEPMSWVPVSSAEKKKNITLDCARGTANCVVFSCLPLYSFDRAAVLHVWGRWNSTFLEEY
SAVKSLEIVIRANTTVKSSIKNMLLRDASTVIPVMVYLDPMMAVVAEGVPVWVILLAVLAGLL
VLALLVLLWKMGGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTTILRNWGSPPRREG
DAHPIAADGHPPELPGDPGHGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCCTGGCT
CACAAACAAGATGCTCAAGGTGTGAGCCGTAAGTGTGTGTGTGTCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGGCGGTGGCTGAGCCGGGGGGCGGTGCGACGGCGGTAATTTTC
TGGATGATAAACAAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGAAC
AAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACCTTGGAGTCCAGGAAAACCCCTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCAGTGGTCTATCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCTCTCAGATAAGCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCGTGAGAGAAGCAGATTGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGTCTTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTGGACAGTCTGGTGTGTTGACAGATATGGAATGAAGTC
ATGGGATCCAGAATAAATGGTGTGTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTATGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAAGTTTATGTTTAAAT
AAGAATCATTGCTTTGAGTTTTTATATTCCTTACACAAAAAGAAAATACATATGCAGTCTA
GTGAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTCACGAGAACAACTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAG
ATAATTCTAAGTGAAATTTAAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAWCQSLSAAAAVAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQDALDPAKDPCCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCFVVPSPFVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKTKTLRPERSRFDTSILPI
CKDSLGMFNRDLTDNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNEWCYCFQ
RQODPPCQTELSNIQKRQGVKKLLGQYIPLCEDGYYKPTQCHGSVGGCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDEDDIMNDEDEIEDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCTGCTCAACACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
CTCTTTTCAGCCCGGGATGCCCCAGACGAGGATGGCGGACAAAGATCTGGCTGCCCTTCCCGTGCTCTTCTGCCC
GCTCTGCCCTCCGGTGCTGCTGCTGGGGGCGCCGGCTTCACACCTTCCCTCGATAGCGATTCACCTTTACCCCTT
CCCGCGGACGAAGAGGAGTGCTTTCTACAGGCCATGCCCCGTAAGGCCCTCGCTGGAGATCGAGTACCAAGTTTATA
GATGGAGCAGGATTATGATATTGATTTCCATCTTGCTCTCCAGAAGGCCAAACCTTAGTTTGTGAACAAAAGAAA
TCAGATGGAGTTTACACCTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTCAGCACCATT
TCTGAGAAGGTTGATTTTCTTGAATTAATCCTGGATATATGGGAGAACAGGCCACAAGAACGAAGAATGGGAAG
AAATATATTCTGGCACAGATATATGGATATGAACCTGGAAGACATCCTGGAATCCATCAACAGCMTCAAGTCC
AGACTAAGCAAAGTGGGCACATACAAATTTCTGCTTAGACCATTTGAAGCTCTGATCGAAGCATCAAGAAAGC
AACTTTGATAGAGTCAATTTCTGGCTCTATGGTTAATTTAGTGGTCATGGTGGTGTCAGCCATTCAAGTTTAT
ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAAGTTAAACCTCCAACTAGAGTACGTAACTTGA
AATGAGGCATAAAAATGCAATAAATCTGTACAGTCAAGACCTTAATGGCTCTTCCAAAAATTTTGGAGATATA
AAAGTAGGAAAACAGGTATATAATTTAATGTGAAAAATTAAGTCTTCACTTTCTGTCGAAGTATCCTGCTGATCCAG
TTGTAATTAAGTGTGAACAGGAATATTTTGCAGAAATATAGGTTTAACTGAATGAAGCCATATTAAATACTGCAT
TTTCTTAACCTTTGAAAAATTTTGCAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCAATTTGCAACACC
AGTCTGTTTAAACAGGTTCTATTACCCGAACCTTTTGTAAATGCGGCAGTTACAATAAATCACTGTGGAAGTTT
TCAGTTTAAAGTTATAAACTCACTGAGGAATTAACCTTAATGATGGATTGAATAAATCTTTAGACTACAAAAAGCCAA
CTTTTCTCTATTATACATATGCATCTCTCTATAATGTAATAAGTAATATAGCTTTGAAATACAAATAGGTTTGTG
AGATTTTATAACAAATACATTTCAAGTGAACATATTAGCAGAAGCAATAGTCTTTGTAATTTGCTTACATTT
CCAAAGCTTGACATTTTTCAGATTCTTAAAAACAAAGTTACACTTACTAAAAATTAGGACATGTTTCTCTTTG
AAATGAAGAATAATGTTTAAAAAGCTTCTCTCCATAGGGACACATTTCTCAACCCCTTAACATAAGTGTAGGA
TTTTAAAAATTAAGTGTGAGGTAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATTCTGCAACAGTTAA
TAATCAATGTTATGTTAATTTTAACTGATTTGCTGACTTGGATAATTCATTATTACAGCAGTTATGAAGGAAATA
TTGCTAAATGATCTGGGCCCTACATAAATAAATATCTCTTTTCTGAGCTCTAAGAATTATCAGAAAAAGGAA
AGAATTTAGAAAACTTGAGAAAACTTAATCCAAATAAAATTCACCTTAAGTAGAATATAAATAATATCTAGA
ATCTGACTGCTCATCATGACCTCTACTCATACATAAATCAAGGAGATGATTAATTTCCAGTTAGCTGCGAAG
AACTTTTGGCTGATAGTTTATTTTCTACAAGAAATCTCGGTTTGAATTTATTTTGAAGCAGGTACATTTTATA
AAATGTAAGCCCTACTGTAGGTTTAGCATATTAACCTACTGCTCTTTTCTATAAAATTTATGATAACAACTTTTAT
TAAAAATGGGCTTTCTGAACTTTATTTATGATGTTGAAGTAAAGGATAGAAACATAGACTCCCAAGTTTAAA
CACCTAAATGTGAATAACCCATATATACAAAGGTTTCTGAGGTGCTCAGGTTGCTATGCTGCAACACTTGCAGAACAA
AGGTCAATAAGATCTTGGCTATGAATACCCCTCCCTTTTGGCGCTGTTAAATTTGCAATGAGAACGAATTTTACA
GTACCATAACTAATAAGCAGGTCAGATATAAATCTGCTCTTTTCTATAAAATTTATGATAAAATTTTAAAGAGG
CCTCTCTGATGCTGTTACTGTACTGCTCTCTGACTCCTTACCTAAACATGAATTTGTTTACATAATCTTCT
ACATGATGATTTGTGCCACTGATCTTAAACCTATGATTCAAGTAACCTCTTACCATATAAAAGCATATTTGCTT
TATTTGGAAGAAATTTAGGAATACTAAGGACAAATTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
CATAACCAAAAAAGCAAACCTGTGAAACAGAGTAAATCTTTAAATTTTCTAAAGCATCTGTTTATCTGCTT
CATATGCTTTTTTAAATTTTCAATTTTCCATTTCTAAATTTAAAGTTATGCTAAATTTGAGTAAAGTGTGTTTACCTT
AACAGCTTTTTTGTCTTTTTTCAATATACAAATTTTAAAAATACATTAATTTTAACTAAGGCCCAACCGATTTC
CATATGTAGCAGTTACCGTGTTCACCTCACAATGAAGCCCTAGAGTTTGTCTGATATGATTTTGGATGATTAAT
GTATGCTGTTCTTTCTCATGTGAATGTCAAGACATGGAGGGTGTGTTGTAATTTTATGTAATAATTTTCTCTTA
CACATAATGGTGTCTTAAATTTGACAAAAATGAGCACTACAATTTGATGTCTCTCTCAATTAAGAGATCTTTAT
TGAAATTTTAAAGACATTTGATTCGTCATGTAAGGATTTTTCATCTGAAGTACAATTAATGACAAATCAAGTGTG
CTCAACTGCTTTATACTTATAACAGCCATCTTAAATAAGCAACGTATTTGTGATCTGATATGATATAAATA
AAATTAACAAAGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPPFVLLLAALPPVLLPGAAGFTPSLSDSFTFTLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFTSTISEKVIFFEL
ILDNMGEQAQEQEDWKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVMVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCCAGTGTGTCTCTGGATG
CTGCTTTCTCTGCCTCATTCTCCTGTGTGAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCATGGCTCCCCCTGCTATGCCTTGT
TTTTGTACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCTCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCATATGTCTGCAAGTTCAAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTGATGATCCTCC
TTCTTTTCTCTTTTCTTACCTTCATTTAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAATGTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIWGLHDPDTPQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTGCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGCTGGCGCTCCTGGCGCTGGTGTGGC
TGCTGTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCGTGTCTGTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAA

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP

FQGDSTVTKSCASKCKPSDVGIGQTLFVSCNTELCNVDGAPALNSLHCGALTLLPLLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTCTCTTTCTTTCTCCCTCTT
GAGTCCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGAGGCGTGCAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAAATGAT
CATAGCACCTTGGATGGGTATTCAGAAGAACCACCTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTGTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTIONAGAAATATTCAGCGTTGTTACTGTGGAGAAGGTCTGTCT
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGA
GACACTAAACAGCTATCCAAATGCAGTGAACTCCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACCTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGCGCACTTACCTGTAAATGCAATGA
AACTTTTAAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTCTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAGCATAACCCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAAAAAAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTCTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTTGGGCGCAGCGGGAGCTAACCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCGGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAACG
CTGCATGCGTCANGCTATGTGCTGCCCGGGAATTACTGCAAAAATGGAATATGTGTGTNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCGCGCGCTGGTAGTCGCGGTTGGCTGCACCTCACCAATCCCGTGC GCGCGG
 CTGGGCGGTGCGGAGATGGCTGTGCTCTCTCTCTGCGACGCGGTGCTTGGGCTCGGCCAGGCGGGTGC GCGCGCA
 AGTGTTCGAGGTAGGGGAGTGCATACAGGAAGCGACCCCGATGGCAAGGTATATTTTGTGGAATGAAAGGA
 AGTATTAGAAATGAGCTGAAGACGATTCACAGATTAAATATTTTGGGAGCAGATTGTGTATGCTTGATTCACCT
 TGAAGTAATGTAGACAGAAGTTCTCAAAATTCGATATACATCAACTGGAACAGCAGTGAATCTTAATGTTCAC
 TTAATTCAGACATTCGCATAGAAGAGAGATGGGAGTCTGGTTAAATAAGATGACTATATCAGAGATCGAAAAG
 GATCATTTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGACAGATCAGGATTTTACAGTTTATCTGG
 AGTGTTCGAAAATCGCAAGCAGTAGAGAAAATAGACAAAGCTTCAAGAAATTTGGCATTGAACTTACATCTGATAA
 AAACCCGGAATACCCAAATGACATGGCGATTTTTAAAAATTAATAGACATATGGAAGTACTCAAGAGTGAAGA
 TCTACCGGAAAAGTATGACAAATATGGAGAAAAGGACCTTGAGGATATCAAGGTGGCAGTATGAAAAGCTGGAA
 CTATTATCGTTATGATTTTGGTATTTATGATGATGATCTGGAATCATAACTTTGGAAAGAGAGAAATTTGATGC
 TGCTGTTAATTCTGGAGAACTGTGGTTTGTAAATTTTACTCCCGAGGCTGTTACACTGCCATGATTTAGCTCC
 CACATGGAGAGACTTTGCTAAAAGAGTGGATGGGTACTTCGAATTTGGAGCTGTTAACTGTGGTGTATGATGAAT
 GCTTTTGGCGAATGAAAGGAGTCAACAGCTATCCAGTCTCTTCAATTTTTCGCTTGGAAATGGCCCGAGTGAATA
 TCAATGGAGACAGATCAAGAGGAGTTTATGTAGTTTTCGAATCGAGCATTTAGAAGTACAGTGACAGAACTTTG
 GACAGGAAATTTTGTCAACTCCATCAAACTGCTTTTGTGCTGGTATTTGGCTGGCTGATCACTTTTGTTCAAA
 AGGAGAGATATTTTGACTCTACAGACACGACTCAGGCTTAGTGGCATGTTTGTTCCTCAACTCATTTGGATGTCAA
 AGAAATATATTTTGAAGTAAATACATAATCTCCAGATTTTGAACACTTTTCGGCAAAACACATAGAGAGTCTGTT
 GGCTCATCATCGGTGGCTGTATTTTTCATTTTGGAAAAATGAAATCTCAAATGATCCTGAGCTGAAAAAATCT
 AAAAATCTACTTAAAAATGATCATATTCAGTTGGCAGGTTTGACTCTCTCGCACAGACATCTGTAGTAA
 TCTGTATGTTTTCAGCCGCTCTGACGAGTATTAAAGGACAAGGAACCAAGAAATATGAAATTCATCATGTGAAA
 GAGATTTCTATAGATATCTTGCTTTGCGAAAGAAAGTGAATTTCTCATGTTTACACGCTTGGACCTCAAAA
 TTTTCTGCGCAATGACAAAGAACATGGCTTGTGTATTTTGGCCCTGGTGTCCACATGTGCGAGTCTTACT
 ACCAGAGTTGCAAGACGATCAAACTCTTCTTATGGTCAGCTTAAGTTGTGATAGTTGTACAGTTCATGA
 GGAATCTGTAACTGTATAAATCTCAGGCTTATCCAACACAGTGGTATTCAACAGTCCAAACATTCATAGATA
 TGAAGACATCACTCTGCGAACTTTCGAGTTTATAGAGGATCTTATGAATCTTCAGTGGTCTCCCTTAC
 GTTCACACCTTCAACGAATAGTTTACAAAGAAAACCAACGAAGCTTGGATGGTGTATTTCTATCTCCGCTG
 CTGCTACCTCTCGAGCTCTTATGCGCAAGTGGAAAAGAAATGGCCCGGACATTAACCTGGAATGATCAAGCTGGG
 CAGTATAGATTGCCAACAGTATCATTTCTTTTGTGCCAGGAAAACGTTCAAAGATACCTCGAGATAAGATTTTT
 TCCCCCAAATCAAATAAGCTTTACAGTATCACAGTTACAATGGTTGGAATAGGATGCTTATTTCTCCGAGAAT
 CTGGGGTCTAGGATTTTACTTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAAAAGTTCTCAAGG
 GAAAAATCTATGGGTGATTTGATTTCTATGCTCTCTTGGTGGGACCTTGCACGAATTTTGGCTCGAATTTGAGCT
 CTGGCTCAGGATGATTAAGGAAAAGTGAAGCTTTTATTTCTACGAAAGCAAGAGAAATTTCAAGAAAGGCA
 GATAAATACAGAGATGCAAAAGCAATCGCTGCCTTAATAGTGAATAATGGAATCTCCGAAATCAAGGCA
 GAGGAATAAGGATGAATCTTGAATATGTTGAAGTAGAAGAAAAGTTTAAAGAAATTTGCAGACGATCAAGCAG
 AAGACACCTATTAGAAATGTACATTTATGATGGGAATGAATGAACATTTCTTACAGTCTCAGTTGTACTGCCA
 GAATTTATCTACAGCAGCTGGTGTAAAGAAAGGCTTGCACAACTTTTCTGTAAGAGGCGCGTTTATAAATTTTTA
 GACTTTGCGAGCTATAATATATGGTTTACACATGAGAACAGAATATAGATCAATGATTTCTTGTATTTTGCT
 TTTAAACACCTTTAAAAAATATTAAAAAGATTTCTAGCTCAGAGGCATACAAAAGTAGGCTGGATTCAGTCCAGT
 GACCATATGATGCTGCTCCCTCGACGGGACTTATAATGTTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
 ATCTACATAAATGTCTAAGTTGTATAAGTCCACTTTCCCTTCAGCTTTTTTGGCTGACCTGAAAAGAGGTAACCT
 TAGTTTTTGGTCACCTTTGCTCTCCCTAAAAATGCTATCCCTAACCATATATTATATTCTGTTTTAAAAACACCAT
 GATGTGGCACAGTAAACAAACCTGTTATGCTGTATTTATGAGGAGATTTCTCATGTTTCTCTTCTCTCTCA
 AAGTTGAAAATAAGCTCTTTATTTTTCACAGCCGAGAACAGTGCAGCAGTATAATGTCACACATGTAAGTACAC
 AAATTTGAGCAACAGTAAAGTGCACAAATCTGTAGTTTGTGTATCATCCAGGAACACTGAGGGAATAAATAAT
 TAGCAATTAATCGGCACTGTAGATATCTCTAAATATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
 TGTGTTCAATGATTTTCTGAAATGCTTTTATAGAAAATTTTCCCACTGATAGTTGATTTTTTGAGGCACTTATAT
 TATCATATTTGCTCTTGAACCTTTGTTGACCTGTATCTTTTATTTAGTATGGGTTTTCTTCTCATGTTTGTG
 TTTTCTACTCTGTCAGCTCATTTTATTTTCAAATAGGAAAAATTACTTTTACAGGTTGTTTTACTGTAGCTTAT
 AATGATACGTAGTTATTTCCAGTTTACTAGTTTACTGTGACAGGCTGCTCTTTTTCAGATAAAATATGACATAATA
 CTCAAGTATTATTTTAAAGAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTGA
 CTTCAAGAGTCAAAATTTTGTCAATGATGTAGTTGTTTATGTTATAATTCAGATGTACAGAATGGTAAAAAT
 CCAATCAGTCAAAGAGGTCATGAATTAAGAGGCTTGAACCTTTTCAAAAAAATAAATAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRD LKRIILCFLIVMAILVGT DQDFYSL LGVSKTASSREIRQAFKKLALKL
HPDKNPNNFNAHGD F LKINRAYEVLKDEDLRKKYDKYGEKGL EDNQGGQYESWNYRYDFGI
YDDDPEIITLERREFDA AVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSVTTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICS NLYVFQP
SLAVFKGQGTKEYEIH HGKKILYDILAFAKESVNSHVTTLG PQNF PANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEHEYEGHHS
AEQILEFIEDLMNPSVVS LTP TTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTD LTPQT FSEKVLQGNHWVIDFYAPWCGPCQNFAP EPELLARMIKGKV KAGKVDC
QAYAQTQC KAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGKRKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGGA
GTCTGTTGGTGAAGTTTTCATTCTCTCAGAGGAGAAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGGTTCTGTGGGATATTAAATAAGCGCGGTGTGGAGGAACTGCAGCTGAGTGCCGAAA
ACTAGGCGCTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCTAGGACATTTTGGATCAGAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCAGAAAGGATTCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCCTT
GGGAAAACTGGTATCAAAACCTCATGTCTCTGCCAGTTTTTGTGAATACTGGGTTACCCA
AAAAATCCAAGCACAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTGAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCCATCGTATATCAATATCTTCTGAG
ACTACAGAAGTTTCTTCTGAAACGCGCCTCAGCGATTTTAAATCGTATGCAGAAATTTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAATGAAATTGAATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTACATTTTT
TCAGTCTGTATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCTCTGTTTCTCAAGAATATTTACGTAGTTTTTCATAGGCTGTGTTTTCTT
TCATGCCCTCTTAAAACTTCTGTGCTTACATAAACATACTTAAAGGTTTTCTTTAAGATAT
TTTATTTTTCCATTTAAGGTTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTGAGCATTTGAAAGATTCCCTAGCCTCTTCTTTT
TCATTAGCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCTCTATATTTCTCCCTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAA'TTTAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAGCTAACACAGGAAAAGGAAGTGTACTGGCTATTACATAAGAAA
CAATGGACCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV
YPADLLSTKDEEITKTTFEVNIGLGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPLYC
SSKFAAVGFGHRLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCAGCAGG
ATGCCCCGGCTCTGCGAGGAAGCTGAAGTGAGAGGCCCGAGAGAGGGCCAGCCCCCGGGGCG
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGTGGTGCTGGGGTCGGTGTTATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCGCGCACTTCTACTTGACACGTCCTTCT
CTAGGCCGCACACGGGGCGCGCGCTGCCCCACGCCGGGCGGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCAGAAAGGAGACGGAGCAGCCGCTGCGCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCAGCGCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCAAGGTGGCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGCAACAACGCCAGCGCGCACTGACCTTCAACAAGTTCTGGCGCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCCTCTTCTGTCGCGACCCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCGGTGCCCATGCTGCGGCTGTACGCCAACACACAGCCTGCCCGCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCCTTGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCTGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGACTGAAAGCTTTTCGCGTTGCTTTTTCTGCGTGCGCTGGAACTGACGCACGCGCACTCC
AGTTTTTTTTATGACCTACGATTTTGGCAATCTGGGCTTCTGTTCCTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTAAGATTAATATATTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVMILLIIVYWDSAGAAHFYLHTSFSRPHTGPPLPTPGPDRDRELT
DSDVDEFLLDKFLSAGVKQSDLPKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRMIVLS
GSLLRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFELENEEFYRKFAVPMRLRYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLEDDAAQLQLLQVDRQLRFPPSYRNR
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

T CGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTCTGTGACCGGGGGCGGGCGCGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCTGGTTTCTGAGACCATCCGCCGATTGGCCGCCTGGATTGTTGTTGTCAACAAC
GCTGGCCACCACCACCCCCACAGAGGCCTGAGGAGACCTCTGCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCACTGGGCCGCATGGGCCAGCCCGCTGAGGTGCGGGCTGCGGCAGTGTTCTTGCCCTCCGA
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGTGCAGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCCCGATATCCCTTCCTTGATTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCACCCCCAAACTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCAGTTAGCAAGGTGCCGGGTACCCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLCDVNVNAGHHPPPQRPEETSAQGFRLLELNLLGTYTLTKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACTG
CTGTTTCTTCTTACCATTTCATCTTCCTGGGGCTGGGCCAGCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCAGCTGAGGAACAGCTCAGAGCTGGCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCAGCC
GTATCCCGTGGACCTGCCGGAGGCACGGTGCTGTGTCTGGGCTGTGTGAACCCCTTACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCCCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTGAAATCACCTGGCCAGAAGCCAGGCCAGCAGCCCCGAGA
CCATCCTCCTTGACCTTTGTGCCAAGAAAGCCTATGAAAAGTAAACACTGACTTTTGAA
GCAAG

100617102501

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPRI PVDLPEARCLCL
GCVNPFTMQEDRSMVSVFVFSQVPVRRRLCPPPPRTGPCRQRAVMETI AVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAACATGGCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCTGGAT
CGCGGCTGTGGCGGCGAGCGGAGGCCCGAGGAGCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTGAGAATGGGAGGCTTTTGCAAAGAA
TGGTGAAATACCTTCAGATCAGTGTGGGGAAGGTAGATGTCACTCAAGAACACAGGTTTGATG
GCCGCTTCTTTGTCACCACTCTCCGACATTTTTCATGCAAAGGATGGGATATTCGCGCGT
TATCGTGGCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCAACGATGTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTACAGTGACTCTTGGAAAT
CCTGCTTGGTGTTCTTATGTGTTTTTCGTATAGCCACCTTGGTTTTTGGCCTTTTTATGGG
TCTGCTCTGGTGGTAATATCAGAATGTTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAAACAAGACAGCCTTGATAGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACCTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCTGCCAGC
TGACACAGAGGTTGGTGAAGACTCCTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGAC
TGTAGATTTAATGATGCGTTTTCAAGAATACACACAAAAAATATGTAGCTTCCCTTTGG
CCTGCAGTTTGTACCAATCCTTAATTTTTCTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAAATCAGGATATAGAAAAACAAACGTAGTGTGGGATCTGTTTGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGGTGAGAGAGTCTCGACCAGAGGAGGCCATTCCAGTCCATAATC
AGCACCTTCAGAGACAAGGCTGCAGGCCCTGTGAATGAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAGGTGAACGTAGAAGCCTTGACATCCTTTCTTGTTGAAGTATTTAT
TTTTGTCAAATTCAGGAAACATCAGGCACCAAGTGCATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTCTTACCCTTAATTTTTCCAGCATTTCCACCATGGGCATTACAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAATGTTGTGGCCATCAGAGACTGAAAGGAAGTAAAGGAT
TTACAAGCAGATTAAAAAAAATTTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATCTGTTTTGTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATACTAAAAAACCACTTCTGATTTTCCCTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACCTCAGTACCTGAAAGTGAAAGATTTGATTTTGTTCATCTTCTGTAATC
TTCCAAAGAATTATATCTTTGTAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMASNWTLVMEGEWMLKFYAP
WCPSCQQTDSWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSESEQNRRSEEAHRAEQLQDAEEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEDNLAAGVDEERSEANDQGPPGEDGVTRE
EVEPEEAEGISEQPCPADTEVVEDSLRQRKSHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSAAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFMSILNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 203

GGAGGGGAGGAGCAGGCCACAGGCCAGGCCGGTGAGGGACCTGCCAGACCTGGAGGGTCTCGCTCTGTCA
CAACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTCAAGTGATTCTCATGCC
TACGCTCCCGAGTAGCTGGAGATTACAGTGGTGACTCCAAAGTAGTACCTCGTCGGAGAGAAATGACTCCCCAG
TCGCTGCTGCAGACGACACTGTTCCTGCTGAGTCTGCTCTTCTGGTCCAAGTGGCCACGGCAGGGGCCACAGG
GAAGACTTTCGCTTCTGCAGCCAGCGGAACGACAGACAGGAGCAGCCTCCACTACAACCCACACAGACCTG
CGCATCTCCATCGAGAACTCCGAAGAGGAGCCCTCACAGTCAATGCCCTTTCCTGCAGCCACCTGCTTCCCGA
TCCCTCCCTGACCCCGAGGCGCTTACCACTTCTGCTCTACTGGAACCGACATGCTGSGAGATTACATCTTCTC
TATGGCAGAGCGTGACTTCTGTGAGTGACAAAGCCTTAGCCTCTCTGTGAGATTCTATGGGAGCATGGTGAGCTC
GCTCAGGGCCCCCGCTGTGTAGCCACTTCTGTCACTCTCTGGTGGAGCCCTCAGAACATCAGCTGCCACGTGCC
GCCAGCTTCACTTCTCTTCCACAGTCTCTCCACACAGGCCCTCATGCTCGGTGGACATGTGCGAGCTC
AAAAGGAGCTCCAGCTGCTCAGCCAGTCTCTGAAGCATCCCGAGAAGGCTCAAGGAGGCGCTCGGCTGCCCC
GCCAGCCAGCATGTCGAGAGCCTGGAGTCGAAATGACCTCTGTGAGATTCTATGGGAGCATGGTGAGCTC
GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCAAGCGGCCCTCCAGACCTGCACATCCACTCCCGG
CAGGAGGAGGAGCAGAGCCGAGATCATGGAGTACTCGGTGCTGCTGCCGCAACACTCTCCAGAGGACGAAGGG
CGGAGCGGGGAGGCTGGAAGAAGACTCTCTCTGGTGGACTCAGCAGCCAGGCCCTGTCCAGGACAAAGAAATTC
AGCCAAGTCTCGGTGAGAAGGCTTGGGGATTGTGGTACAGAAACCAAAAGTAGCCAACTCCAGGAGCCCGTG
GTGCTCACTTCTGCAGCCAGGACTCAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTGAAGACCCCA
TTGAGCAGCCCGGGCATTTGAGAGAGTGTGGTGTGAGACCGTCAGGAGAGAAACCAAACTCTGCTTCTGCG
AACAACCTTGACCTACTTTGACGTCTGATGCTCTCTCGGTGGAGGTGGAGCGCGTGCACAAAGCACTGCTGAGC
CTCTCTCTACGTGGGCTGTGCTGTCTGCGCTGCGCTGCTTGTACCATTTGCCGCTTACCTACTCTGCTCAGG
GTGCCCTCGCTGCGCAGGAGAAACCTCGGAGCTACACCATCAAGGTGCACATGAACCTGTGCTGGCGCTCTTC
CTGCTGGAACAGGACTTCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTGTGAGGCTGCTGCTCGGAGCCAGTGCC
ATCTTCTGCACTTCTCTGCTCACTGCTCTTCTGGATGGGCTCGAGGGGTACAACCTCTACCGACTCGTG
GTGGAGGTCTTTGGCACTTATGCTCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGCTTCCCATCTTT
CTGTGTACGCTGTGGTGGCTTGTGGATGTGGACAACATAGGCCCATCATCTTGGCTGTGCTAGGACTCCAGAG
GGCGTCACTTACCTTCTGCTGGATCCGGGACTCCCTGGTCAGTACATCAACCACTGGGCTTCTTACG
CTGTGTGTTCTGTGTCAACATGGCCATGCTAGCCACCATGTGTGTGCAGATCTCGGGCTGGGCCCCACACCCAA
AAGTGTGCACATGTGCTGCACTGCTGCGGCTCAGCCTGCTCTGGGCTGCGCTGCTGCTTCTTCTCTTCT
TTGCTCTTGCCACTCTCAAGCTTGTGCTCTTCACTTTTACGATCATCACTCTCTCCAAAGGCTTCTCATC
TCTACTGTGACTGTGCTCATCGGCTCAGGCCCGGGTGGCCCTCCCTCTGAGAGCACTCAGACAGGCC
AGGCTCCCATCAGCTCGGCGACACCTGCTCAGCCGCATCTAGGCCTCAGGCCCACTGGCCATGTGATGAAG
CAGAGATGCGGCTGCTGCACTGCTGTGCCCCGAGCCAGCCAGCCAGCCAGCTCAGCCGAGACT
TTGGAAGCCCAACGACCATGGAGAGATGGCGGTGGCATGTGTGAGCGGACTCCCGGCTGGCTTTGAAATG
GCCTTGGGAGTACTCGGCTCTCATCAGCTCCACGGGACTCAGAAGTGGCGGCTGCTGCTAGGATCTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCTGTGCTCTCTTCAACCCCTCGGCGCCAGCCCTCATTTGCTGG
GGCCAGGCCCTTGATCTTGGGGTCTGGCACTCTTAATCTGTGCCCTGCTGGGACAAATGTGCTGCTCA
GTTGCTCTGCTCTCTGTGTCACCTGAGGGCACTCTGCATCTCTGTCAATTTAACTCAGGTGGCAACCCAGG
CGAATGGGGGCCAGGCGAGCCCTCAGGGCCAGAGCCTGGCGGAGGAGGCGCCTTTGCCAGGAGCAGAGCAG
AGCTCGCCTACTCTGAGGCCAGGCCCTCTCTCTCTCAGCCGCCAGTCTCTCTTCACTTCTCTCGGGTTC
TCTCTCTCTCCAGGGCGCTTCTGCTCTCTGTTCACTGCTGGGGTCCCGGATTTCAATGCTGTTTTTGGGGA
GTGGTTCTCAGAGAGTCTCGTGTGCTCTGCTGTAATGTTGTGCTACTGCAAGCCTCGGCTCGGCTCGGCA
GGCTCGGTACCGATGCTGGGCTGGGCTAGGTCCTCTGTGCTCTGGGCTTTGTATGAGCTGCATTTGCCCTTG
CTCACCTACCAAGCAACAGCTCAGAGGGCGCTCAGCCTCTCTGAGAGCCCTCTTGTGGCAAGAACTGTGGA
CCATGCGAGTCCGCTGTGTTCCATCCCACTCTCAAGGAGTGAAGTGAACCTCTGTGTGACACTGGGCTCA
GAGCTCGCACTCTCTAAGAGGTCTCTTCAAGCCCAATAAGTCTCAGGCGCCTCGGCCCCCATCATGTT
TAATCTCTGCTCAACAAACACACCGGGTAGATTGCTGGCCTGTTGTAGGTGTGAGGACACAGATGACCGACCTG
GTAATCTCTCTGCAACATCAGTCTGATGTGAGGCGTGGTGAAGCAAGACTCTCGGAGACTCAGGGACA
GGGAGCATCATCTGCTGCTGGGAATCTGGAAGACTCTCTGAGGAGTCAAGCGTTCATCTTGACCTTGAAGAT
GGGAAGGATGTTCTTTTACGTACCAATCTTTTGTCTTTGATATAAAAAGAGATACATGTTCTATTGTAGAGA
ATTTGGAACCTGTAGAAGAGAAATCAAGAGAGAAAAATAAAATCAGCTGTTGTAATCGCTAGCAAAAAA
AAAAA

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MT PQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFC SQRNQTHRSSLHYKPTPDLRISIENSE
EALT VHAPFPAAH PASRSFPDPRGLYHFC LYWNRHAGRLHLLY GKRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNI SLPSAASF TFSFHPHTAAHNASVDMCELKRDQLL
SQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLT FQHQLQPKNVTLCQVFVVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVD AVHKHYLSLLSYVGCVVSA LACLV TIAAYLCSRVPLPC
RRKPRDYTIKVHMNLLLA VFLLDTSFLLSEPVALTGSEAGCRASAI FLHFSLLTCLSWMGLE
GYNLYRLVVEVF GTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGP IILAVHRTPEGVIY
PSMCWIRDSLVS YITNLGLFSLVFLFNMA MLATMVVQILRLRPHTQKQWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLVLYFSIITS FQGFLIFIWYSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

[illegible]

Table 1. Demographic characteristics of study population

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTTCAGGTCCAGGTTTTCGCTTTGA
TCCCTTTTCAAAAACCTGGAGACACAGAAAGGGCTCTAGAGAAAAGTTTGGATGGGGAATATGTGGAACACACCT
GGGATTTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCCACTGCGAGCTTCCCTGGCGGTGGTGAAGAGAC
TCGGAGGTCGCTGCTCCAAAAGTCCGCGCTGAGTGAGCTCTCACCCCACTCAGCCAAATAGAGCCTCTTCGGGG
TTCTCTCTGACATCTGCCTTGGCCGCGGACGAGACAGGGGACTCAGGCGGAATCAACCTTGAGTAGATAAATTC
AGTTTTCAGACACAGGAAACAGAACGGAGTACAAGATCCCTCAGCATGAGAGAAATTTATCTGTGTCTACTAATG
GAGATATTCACAGCCCAAGGTTTCTCATACTTATCCAAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
GAAAGAAATGTATGGATACAACCTTACGTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGCATATGCCAAT
ATGATTTGTAGAGTTGAGGAAACCCAGTGATGAACTATATAGGCGCTGGTGTGGTTCTGTACTGTACCAG
GAAACACAGATTTCTAAAGGAAATCAATTAGGATAAGATTTGTATCTGATGAATATTTCTCTTGAACCAAGGGT
TCTGCATCCACTACAACATTTGTCATGCCAACATTCACAGAAAGCTGTGAGTCCCTCAGTGCTACCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAAATGCTTATAGTCCCTTAGTACCTTGAAGAGACCTTATCGATATCTTGAACCA
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTCTTGGCAAGGCTTTTGTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAAGATTACAGCTGCACACCTCGTAACTTCT
CAGTGTCCATAAGGGAAGAATCAAAGAGAACCGATACCATTTCTGGCCAGGTTGTCTCTGGTTAAACGCTGTG
GTGGGAACTGTGCTGTTGTCTCCACAAATGCAATGAATGTCAATGTGTCCCAAGCAAAAGTTACTAAAAAATACC
ACGAGGTCCCTCAGTTGAGACCAAGACCGGTGTGAGGGGATGCACAAATCACTACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGACAGAGGAGCACAGGAGGATAGCCGATCACCACAGCAGCTCTTGCCCA
GAGCTGTGAGTGCAGTGGCTGATTCTATTAGAGAACGATGCGTTATCTCCATCCCTTAATCTCAGTTGTTTGCT
TCAAGGACCTTTTCACTTTCAGGATTTACAGTGCAATCTGAAAGAGGAGACATCAACAGAAATTAGGAGTTGTGCA
ACAGCTCTTTGAGAGGAGGCCATAAGACAGGAGAAAGGCTTCAATCGTGGAAAGAAATTAATGTGTAT
TAAATAGATCAGCAGCTAGTTTCAGAGTTACCATGTACGTATTCACCTAGCTGGGTTCTGTATTTTCAGTCTTTT
GATACGGCTTAGGGTAACTGTGTCAGTACAGGAAAAAAGCTGTGCAAGTGAGCACCTGATCCGTTGCCTTGCTTAAC
TCTAAAGCTCCATGTCTCGGGCTTAAAGATCGTATAAATCTGGATTTTTTTTTTTTTTTTGTCTCATATTACAT
ATGTAAACACAGAACTTCTATGTACTACAACTGGTTTTTAAAGGAACTATGTGTGTATGAATTAACCTGT
GTCATGCTGATAGGACAGACTGGATTTTCATATTTCTTATAAATTTCTGCCATTTAGAAGAGAGAACTACA
TTCATGTTTGAAGAGATAAACCTGAAAGAGAGAGTGGCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGACATTTTTATATCTCTCTTTGACATTAACTGTTGGCTTTTCTAACTGTAAATATATCT
ATTTTTACCAAAGTATTTAATATCTTTTTTATGACAACTTAGATCAACTATTTTGTAGCTTGGTAAATTTTTCT
AAACACAATTTGATAGCCAGAGGAAACAAAGATGATATAAATAATTTGTCTTCGACAAATAACATGATTTTCA
TTCTGTATGGTGCTAGAGTAGATTAATCTGCATTTTAAAAAAGTAATTGAATAAGAAATTTGTAAGTTGCAAA
GACTTTTTGAAAATAATTAATATCATATCTTCAATTCCTGTTATTTGGATGAAATAAAAGCAACTTATGA
AAGTAGACATTCAGATCAGCCATTACTAACCTTATCTTTTTTGGGGAATCTGAGCTAGCTCAGAAAACAT
AAAGCACCTTGAAGAAAGACTTGGCAGCTTCTGATAAAGCGTGCTGTGCTGTGCATACATGTATAAATACATGATTTA
TTGTGATGTGTGGTTTTATATCTTAAACCTGTCCATACACTGTATATAACATGAGATTTTTATGTACA
GAAGTATGTCTCTTAACCACTGCTATTTGTACTCTGGCAATTTAAAGAAAAATCAGTAAAAATTTTTGCTGT
AAAAATGCTTAATATNGTGCTAGGTTATGTGGTGACTTTTGAATCAAAAAATTTGAATCATCAAAATAAAGA
ATGTGGCTATTTTGGGAGAAAATTAAGAAAAAAGGTTTAGGGATAACAGGTAATCGCGCC

FIGURE 207

MSLFGLLLLSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGKRSRVVDLNLLEEVRLY
SCTPRNFSVSIREELKRDTIIFWPGCLLVKRCGGNCACCLHNCNECQCVP SKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDVCVRGSTGG

Signal sequence:

amino acids 1-14

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2

FIGURE 209

MVFPMTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIDCTDKHLTEIPGG
IPTNTNLTNLTLTINHIPDISPASFHRLDHLVEIDFRNCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIKENLTLANIEILYLGNQNCYR
NPCYVSYSIEKDAFLNLTCLKVLSLKDNNVTAVPTVLPSTLTLEYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILIRGYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYQQ
TLDLSKNSIFFVKSSDFQHLNFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVQLKLMNDNDISSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLELDISKNSLSFLPSGVFDGMPPNKLNLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNMLKMLLLHHNRFLCTCDAVWFVWVWNHTEVTIP
YLATDVTCVGPAGHKQGSVISLDLYTCELDLTNLILFSLSSISVSLFLMVMMTASHLYFDVWV
YIYHFCKAKIKGYQLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPQGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIIIFLE
KPFQKSFLQLRKRLCGSSVLEWPTNPQAHFYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTTGGCGTGTGCAAGTTACGGAATGAAAAATTAGAACACAGAAACATGGAACAACTGTTCTCTTCT
 AGTCGTCAATGCTGACCTGCATTTTCTCTGCTAATATCTGGTTCCTGTGAGTTATGCGCGGAAGAAAAATTTTCTA
 GAAGCTATCTCTGTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGACAGGTGCAGCAATCGTGACTACAGGAAG
 TTCCCAACCGGGTGGGCAAAATATGTGACAGAACTAGACCTGTCTGATAATTTTCATCACACACATAACGAATGAAT
 CATTTTCAAGGGCTGCAAAATCTCACTAAAAATAATCTAAACACCAACCCCAATGTACAGCACAGCAAGCGAAATTC
 CGGGTATACAAATCAAAATGGCTTGAAATATCACAGACGGGGCAATCTCTCAACCTAAAAAACCTTAGGGAGTTACTGC
 TTGAGAGCAACACAGTTTACCCCAAAATCCCTCTGGTTTGGCAGAGTCTTTGACAGAACTTTAGCTCTAATTCAAAA
 ATATATATACAAATCAATAGCTAAGAGGGGCATTTCAGAGCTATATAAATCTGAAAAATCTCTATTTGGCTGGAACTGCT
 ATTTTAAACAAGATTTGCGAGAAACATAACATAGAGAGATGGAGTATTGAAACGGCTGACAAATTTGGAGTTGCTAT
 CACTATCTTTCAATTTCTCTTTCAACAGTGGCCACCAACCTGCCAAGCTCCCTACGCAAACTTTTCTGAGCAAC
 CCCAGATCAAAATACATTTAGTGAAGAAGATTTCAAGGGATGTATAAATTTAACTACTAGATTTAAGCGGGAATCT
 GTCCGAGGTGCTTCAATGCCCATTTTCCATGCGTGCCTTGTGATGGTGGTCTTCAATTAATATAGATCGTTTGTG
 CTTTTCAAAACCTTGACCCAACTTGATACCTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGT
 TTAATAATATGCTCATCTGAAGGTGCTGGATCTTGAATTCAACTATTAGTGGGAGAAATAGTCTCTGGGGCAT
 TTTTAAACGATGCTGCCCGCTTAGAAAATCTTGACTGTCTTTAACTATATAAAGGGGAGTTATCCACAGCATA
 TTAATATTTCAGAAACCTTCTCTAAACTTTTGTCTCTACGGGCATTTGCATTTAAGAGGTTATGTGTTCCAGGAAC
 CTAGAGAGATGATTTCCAGCCCTGATGCACTTCTCAAATTTATCGACTATCACTTTGGGTATTAATTTTATTA
 AGCAAAATGTTTCAAATCTTTCAAATTTCTCAAATCTGGAAATTTATTTGTGACAGAAACAGCAATATCAC
 CGTGTGTTAAAGATTAACCGGAGAGTTATGCAAAATAGTTCCTCTTTCAACGTCTATCTCGGAAACGACGCTCAA
 CCGATTTTAAAGTTTGACCCCAATTCGAACTTTTATCATTTTCAACCGTCTTTTAAATAGCCCAATGTGCTGCT
 ATGGCAAGGCTCTAGATTTAAGCGCTCAACAGTATTTTCTTCAATGGGCGAAACCAATTTGAAAAATCTCTGACA
 TTGGAATTTTAAATCTGTCTGCAAAATAGCAATGCTCAAGTGTAAAGTGAATCTGAAATTTTCAGCACTTCTCATG
 TCAAAATTTTGGATTTGACAAACAATAGACTAGACTTTGATATGCTAGTGTCTTACTGAAATTTGCTGCAGCTGG
 AAGTTCTAGATCTCAGCTATAATTCACATAITTTAGAAATAGCAGCGTACACATCATCTAGAAATTTTAAACA
 ATTTTCAAACTCAAAGTTTTTAAACTTGAGCCACAACATTTATACTTTAACAGATAAGTATAACCTGGAAA
 GCAAGTCCCGTGAAGATTTAGTTTTCAGTGGCAATCGCCTTGACATTTTGTGGAATGTAGTGAACAGGTATA
 TCTCCATTTTCAAAGGTCTCAAGAACTTGACAGCTCTGGATTTATCCCTTAAATAGGCTGAAGCACATCCCAAAT
 AAGCATCTCTTAATTTGGCAGCACTCTCACTCAACTACATATAAATGATAATTTTAAAGTTTAACTGGA
 CATTACTCCAGCAGTTTCTCTGCTCTGAGTTGCTTGACTTACGTGGAAACAACTACTCTTTTAACTGACTAGCC
 TATCTGAGTTTACATCTTCCCTTCGGACCTGCTGCTGAGTCTAATGACAGATTTCCACCTTCACTTGGCTTTC
 TTTCTGAAGTCTAGTCTGGAAGCACTCGATTTAAGTTCCAATCTGCTAAAAACAATCAACAATCCGACCTTG
 AAACCTAAGACCAACCAAAATTAATCTAGTTTGGAACTACACGAAACCTTTGAATGCACTGCACTGTGACATTTGGAG
 ATTTCCGAAGATGGATGGATGAACATCTGAATGCTCAAATTTCCAGACTGCTAGATGTCAATTTGTCGCACTGCTG
 GGGATCAAGAGGGAGAGATTTTGTAGTCTGAGCTAAACACTTGTGTTTCCACATCCCAATTTTCTAGTGTGTT
 TTTTCTCACGTTCTTTATCACCAATGTTTATGTTGGCTGCCCTGCTCACCATTGTTTACTGAGGTGTT
 GGTTTATATATAATGTGTGTTTAGTCAAGGTAAGGCTACAGGCTCTTTTCCACATCCCAATTTTCTATGATG
 CTACATTTCTTTATGACCAACAAGATGCTCTGTTTACTGACTGGGTGATAAATGAGCTGCGTACCACCTTGAAG
 AGAGCGGAGACAAAAAGTTCTCCTTTGCTAGAGGAGAGGGATTTGGACCCGGGATTTGGCCATCTAGCAAC
 TCATGACAGACATCAACCAAGACAGAAAAAGTATTTGTTTAAACCAAAAAATATGCAAAAGCTGGAACCTTA
 AAACAGCTCTTTTACTTGGCTTTGACAGGCTAATGGATGAGAACTGGATGTGATTTATTTCTCTGCTGGAGC
 CAGTGTTCACAGATCTCAGATAITTTGAGGCTACGGCAGCGGATCTGTAAGAGCTTCATCTTCCAGTGGCTGACA
 ACCGAAGGACAGAGGCTGTTTGTGGCAACTCTGAGAAATGTGGTCTTGACTGAAATGATTCAGGGTATAACA
 ATATGATGTGATTCATTAAGCAATACTAACTGACATTTTTCGGCCATTAATGAGAGTCAACAGTGAACAA
 GAATGACATTTCTGATAGTATCTATTGCTATGTAACAAATTTATCCCAAACTTAGTGTTTTAAAGCAACACA
 TTTGCTGGCCACAGTTTTTGGAGGTCAGGAGCTCAGGCCAGCATAAATGGGCTCTGCTCAGGTTGCTTCAG
 AGGCTGCAATGTAGTGTCTCACCAGAGACATAGGCATCACTGGGGTCACTCATGTGTTGTTTCTGGAAT
 ATCTCTCTGGGCTATTGGCCAAAGGCTATACATCATGTAAGCATGCGAGCTCTCCCAACAGGCACTGTCTTC
 ATCAGAGCTAGCAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTGTAATCGAATCAAAAAGTAT
 ATCTCATCTTTGGCCATTTCTATTTGTAAGTAAACCCAGGCTCCACAGCTCAGTGGAGTGACCACT
 CTAGTCTCAGGAAACAGCTGGAAGCAAGATGGTGAAGTCTGATGCTTCAAGTGGTCACTCAATATTTTCCCT
 TGACTGCTGTCTTAAAGTGGCTGCTATCTTGATGATAGATTGGAATTCAGTGTGGTGTGCTGCTGGTGTGTTAT
 TTACAGTTGCTTTTAAACAAATTTGCTGTAACATTTGACTTCAAGGTTTATGATGTTTAAAGAACTGAGATG
 ATAGCTTTTAAAGCATCTTTACTCTTACCATTTTAAAGTATGACAGTAAATTTGAAGCTTTTGGTCTATA
 TTGTTAATGGCATTGCTGTAAATCTTAAATGAATGAATAAAAATGTTTCATTTTCAAAAAA

FIGURE 211

MENMFLQSSMLTCTIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVPTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLHNPNVQHQNPNQIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSTLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLLELLSLSFNLSLHVPPKLPSSLRKLFSLNTQIKYISEEDFKGLINL
TLDDLSGNCPRCFNAPPCVPCDGGASINIDRFQNLTLQRLYNLSSTSLRKINAAWFKNM
PHLKVLDLEFNLYLVEIVSGAFLTMLPRLEILDLSFNLIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFSLNLEIYYLSENRIPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVXYLDLTNNRLDFDNASALTELSDELV
LDLSYNSHYFRIAGVTHHLEFIQNFITNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYIISIFKGLKNLTRLDLNLRLKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLTLSSHNRISHLPSGFLSEVSSLKHLDL
SNLLKTINKSALETKTTTKLSMLELHGNPFECTCDIGDFRRWMEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFFFITTMVMLAALAHHLFYWDVWFYINVCLAKVK
GYRSLSTSQTFFDAYISYDTKASVTDWVINELELYHLEESRDKNVLLCLEERDWDPLAIIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLALQRLMDENMDVIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLFWQTLRNVVLTENDSRYNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCAGGACCGGGGAGGCACAGTGGCCCCCACCACCCGAGGGA
GCAGCTCCTGCCCTGTCCGGGGGATGACTGATTCTCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCTGGAGGCA CAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCTGTCTCCGAGTCGTTCTGTGCACGCTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGCCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCTTCCAGGCCCTCGCTACGCGTGTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCTTGGGGCTGTGGAGCAGCAATATGCCAGCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGTGCCGTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGTCTAGGAGGGGCGGCTGTCCCAGCGCTGCATCAACACCGCCG
CAGTTACTGTGTGCCAGTGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGTCTGGAGGAGAAGCTGCAGCTGGTGTGCG
CCCACCTGCACAGCCTGGCCTCGCAGGCACCTGGAGCATGGGCTCCCGGACCCCGCAGCCTCC
TGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCTCTG
GAGGAGCAGCTGGGGTCTTGTCTCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTG
GACTGAGCCCTCAGCCGCCCTGCAGCCCCCATGCCCTTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCTCTTCTCTCTCCCC
TTCTCTGGGAGGCTCCCCAGACCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGTACCCCCACCTGGCTACCCCAACGGCATCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCA CAATAAAAAATGAAACGTGA
AAGGGCGGCCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTATTATGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLA VGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSL SADGTLCV PKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCGTGTGGTTGGCAGCAAAGTTACAGCTTGGCTGGGCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGGCCCGGGAGGGGAAC TGCCCGAGGGAGAGGAAACCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCGGAGGAGCAGCTCCTGCCCTGTCCGGGGGATGACTGATTTC
TCCTCCGCCAGGCCACCAGAGGAGAAGGCCACCCCGCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGACCTGTCTCCGAGTCGTTCCG
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGAGCCCTGGGCTGGCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGTGCGCGTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG
TCCCCAGCGCTGCATCAACCCGCCGCGCAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCCTTCTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGA
CTCGTGA CTGCCAGCGCTCCAGGCTGGA CTGAGCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCTCCTCTTCTCCTCCTCCCTTCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCTGGCTACCCCAACGGCA
TCCCAAGGCCAGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGGCCTCAGTGGGGCTGCTGCCTGAC
CCCCAGCAATAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
21

CCACACGGCTCGGAAGCTGGCCCTGCACGGCTGCAAGGAGGAGCTCTGTGGACAGGCCAGGCCA
GGTGGGCCCTCAGGAGGTGCTCCAGGCGGCCAGTGGGCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCG
CTGGAGGCACAGGCCATGAGGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCAGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCAGC
GGGACCCCTGTCTCCGAGTCGTTCTGTGCAGCGTGTGTACCAGCCCTTCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGACAGCCCTGG
GCTGGCCCTGCCAGGCCTCGCTACGCGTGCTGCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCTGTGGAGCAGCAATATGCCAGCCGCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTTGGGAGGGGCACAGCCCTGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGG
CCCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA
GCCTGGCCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCCTCC
TTCCAGCAGCTCGGCCGATCGACTCCCTGAGCGAGCAGATTTCTTCTTGAGGAGCAGCT
GGGGTCTCTGCTCTGCAAGAAAGACTCGTGGACTGCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCTGCCAACATGCTGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCTCTTCTCTCTCCCTCTCGGGAG
GCTCCCCAGACCTTGGCATGGGATGGGCTGGGATCTTCTCTGTAATCCCTCTGGCTACC
CCCACCTGGCTACCCCAACGGCATCCCAAGGCAGGTGGGGCTCCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCTGGGACCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAAATAAAATGAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDVSESFVQRVYQPFLLTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCVNTAGSYWCQWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
219

FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCTTGTTCAGGAGGAGACAGCCTCCGGCCCCGGGGAG
GACAAGTCGCTGCCACCTTTTGCTCGCGACGTCGATTCCTGGGACGGTCCGTTTCTCGCGTCAGCTGCCGGCCG
AGTTGGGTCTCGGTGTTTCAGGGCCGGCTCCCCCTTCCTGGTCTCCCTTCTCCGCTGGGCCGGTTTATCGGGAGG
AGATTGTCTTCCAGGGCTAGCAATTTGACCTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCTCTGTTTCTCTCTCGTGTAAATCGCAAAACCCATTTTGGAGCAGGAATTTCAATCA
TGTCTGTGATGGTGGTGAGAAAGAGGTGACACGGAATGGGAGAACTCCACGGCAGGAACACCTTTTGCTGTG
ATGGCCGCTCATGATGGCCCGGCAAAAGGGCATTTTCTACCTGACCCCTTTTCTCATCTCGGGACATGTACAC
TCTTCTTCGCTTTTGATGCGCTACCTGGCTGTCAGCTGTCTCTGCGCATCCCTGTAATTTGCTGCCATGCTCT
TCCTTTTCTCCATGGCTACACTGTTGAGGACCGCTTCAGTGACCTGGAGTGATTCCTCGGGCGCTACACAGATG
AAGCAGCTTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGGGTGGCCCCAGGGCCAGCGACCAACCGCTCGTA
TCAAAGATTTCCAGATAAACAACAGATTGTGAACCTGAAATACTGTTACACATGCAAGATCTTCGCGCTCCCC
GGGCTCCCATTTGCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCTGGGTGGGGAATTGTG
TTGGAAGAGGAACTACCGCTACTTCTACCTCTTCATCCTTTCTCTCTCCCTCTCACAACTCATGTCTTCGCT
TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAAATTGGCTTCTTGGAGACATTGAAAGAACTCCTGGAACTG
TCTAGAACTCTCATTTGCTTCTTACACTCTGGTCCGTCGTGGGACTGACTGGATTTTCATCTTCTCTCGTGG
CTCTCAACAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGGCTCCAGAATCCCTACAGCC
ATGGCAATATTTGAAGAACTGCTGTGAAGTGTGTGTGGCCCCCTTGCCCCCAGTGTCTGGATCGAAGGGGTA
TTTTGGCACTGGAGGAAAGTGAAGTGCACCTCCAGTACTCAAGAGACAGTAGCAGCTTTCGCCACAGAGCC
CAGCCCCCAGAGACCACTGAACTCAAATGAGATGCCGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCAAGGAGGCGAGTGAAGCTGAGAAGTAGCCCTATCTATGGAAGAGACTTTTGTTTGTGTT
TAATTAGGGCTATGAGAGATTTTCAAGTGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
GTTTTTCTTTGGTCTTTTAGTACCCAGTTGCACACTGGCATTTTCTGCTGCAAGCTTTTTTAAATTTCTGAACT
CAGGCGAGTGGCAGAGATGTCACTCACTGTATACTGGAATAATGGGTCTCTTGGGCCCTGGCACTGGTCTCT
CCATGGCTCAGCACCAGGGTCCCCCTGGACCCCTCTCTCTCCCTCCAGATCCAGCCCTCTGCTTTGGGGTCA
TGGTCTCATTTCTGGGGCTTAAAGTTTGTGAGACTGGCTCAAATCTCCCAAGCTGTGCACTGTCTGAGTCCAGA
GGCAGTCAAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAAGAGATTCTCGGGCCAACAACTGCGCAGCATGCCCCAAATCTTTTAGGAATGGGACAGGTACCT
TCCACTTGTGTANNN
CAGGAATGGCAGTAATAAAAGTCTGCACCTTTGGTCTATTTCTTTCTCAGAGGAGCCGAGTGTCTCACTTAAAC
ACTATCCCCCTCAGACTCCCTGTGTGAGGCTTCAGAGGCCCTGAATGCAAAATGGGAAACCAAGGCACAGAGAG
GCTCTCTCTCTCTCTCTCTCCCCCGATGTAACCTCAAAAAAATAAGTCTAAACCACTTCTCCATTAAAGCT
CGGCTGAGTGAGGGAAAGCCGAGCACTGTGCCCCCTCGGGTAACTCACCTAAGGCTCTCGGCCACTCTGGCT
ATGGTAACCACTGGGGGCTTCTCCAAAGCCCCCTCTCCAGCACTTCCACCGGCAGAGTCCACAGGCACTT
CACCTGGGGGTGGGCTGTGGCCCCAGTCAGCTCTGCTCAGGACCTGTCTATTTCAGGGAAGAGATTTATGT
ATTATATGTGGCTATATTTCTAGAGCACCTGTGTTTCTCTCTTCAAGCCAGGCTGTCTCTGGATGACTTAT
GGGTGGGGAGGTAAACCGGAACCTTTTCATCTATTGAAGGCGATTAACTGTGCTAATGCA

10016177.102501

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTFLILGTCTLFFAFECRYLAV
QLSPAIPVF AAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHC PWVGN CVGKRNYRYFYL
FILSL SLLTIYVFAFNIVYVALKSLKIGFLET LKETPGTVLEV LICFFTLWSV VGLTG FHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVL CGPLPPSVLD RRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPEPPQEAAEA EK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCCTGATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCCTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCCTTCT
CCCACAGAGCNCCTTCGACCATCACTGCCCTGGGTGGGAATTGTGTTGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

1234567891011121314151617181920212223242526272829303132333435363738394041424344454647484950515253545556575859606162636465666768697071727374757677787980818283848586878889909192939495969798991001011021031041051061071081091101111121131141151161171181191201211221231241251261271281291301311321331341351361371381391401411421431441451461471481491501511521531541551561571581591601611621631641651661671681691701711721731741751761771781791801811821831841851861871881891901911921931941951961971981992002012022032042052062072082092102112122132142152162172182192202212222232242252262272282292302312322332342352362372382392402412422432442452462472482492502512522532542552562572582592602612622632642652662672682692702712722732742752762772782792802812822832842852862872882892902912922932942952962972982993003013023033043053063073083093103113123133143153163173183193203213223233243253263273283293303313323333343353363373383393403413423433443453463473483493503513523533543553563573583593603613623633643653663673683693703713723733743753763773783793803813823833843853863873883893903913923933943953963973983994004014024034044054064074084094104114124134144154164174184194204214224234244254264274284294304314324334344354364374384394404414424434444454464474484494504514524534544554564574584594604614624634644654664674684694704714724734744754764774784794804814824834844854864874884894904914924934944954964974984995005015025035045055065075085095105115125135145155165175185195205215225235245255265275285295305315325335345355365375385395405415425435445455465475485495505515525535545555565575585595605615625635645655665675685695705715725735745755765775785795805815825835845855865875885895905915925935945955965975985996006016026036046056066076086096106116126136146156166176186196206216226236246256266276286296306316326336346356366376386396406416426436446456466476486496506516526536546556566576586596606616626636646656666676686696706716726736746756766776786796806816826836846856866876886896906916926936946956966976986997007017027037047057067077087097107117127137147157167177187197207217227237247257267277287297307317327337347357367377387397407417427437447457467477487497507517527537547557567577587597607617627637647657667677687697707717727737747757767777787797807817827837847857867877887897907917927937947957967977987998008018028038048058068078088098108118128138148158168178188198208218228238248258268278288298308318328338348358368378388398408418428438448458468478488498508518528538548558568578588598608618628638648658668678688698708718728738748758768778788798808818828838848858868878888898908918928938948958968978988999009019029039049059069079089099109119129139149159169179189199209219229239249259269279289299309319329339349359369379389399409419429439449459469479489499509519529539549559569579589599609619629639649659669679689699709719729739749759769779789799809819829839849859869879889899909919929939949959969979989991000100110021003100410051006100710081009101010111012101310141015101610171018101910201021102210231024102510261027102810291030103110321033103410351036103710381039104010411042104310441045104610471048104910501051105210531054105510561057105810591060106110621063106410651066106710681069107010711072107310741075107610771078107910801081108210831084108510861087108810891090109110921093109410951096109710981099110011011102110311041105110611071108110911101111111211131114111511161117111811191120112111221123112411251126112711281129113011311132113311341135113611371138113911401141114211431144114511461147114811491150115111521153115411551156115711581159116011611162116311641165116611671168116911701171117211731174117511761177117811791180118111821183118411851186118711881189119011911192119311941195119611971198119912001201120212031204120512061207120812091210121112121213121412151216121712181219122012211222122312241225122612271228122912301231123212331234123512361237123812391240124112421243124412451246124712481249125012511252125312541255125612571258125912601261126212631264126512661267126812691270127112721273127412751276127712781279128012811282128312841285128612871288128912901291129212931294129512961297129812991300130113021303130413051306130713081309131013111312131313141315131613171318131913201321132213231324132513261327132813291330133113321333133413351336133713381339134013411342134313441345134613471348134913501351135213531354135513561357135813591360136113621363136413651366136713681369137013711372137313741375137613771378137913801381138213831384138513861387138813891390139113921393139413951396139713981399140014011402140314041405140614071408140914101411141214131414141514161417141814191420142114221423142414251426142714281429143014311432143314341435143614371438143914401441144214431444144514461447144814491450145114521453145414551456145714581459146014611462146314641465146614671468146914701471147214731474147514761477147814791480148114821483148414851486148714881489149014911492149314941495149614971498149915001501150215031504150515061507150815091510151115121513151415151516151715181519152015211522152315241525152615271528152915301531153215331534153515361537153815391540154115421543154415451546154715481549155015511552155315541555155615571558155915601561156215631564156515661567156815691570157115721573157415751576157715781579158015811582158315841585158615871588158915901591159215931594159515961597159815991600160116021603160416051606160716081609161016111612161316141615161616171618161916201621162216231624162516261627162816291630163116321633163416351636163716381639164016411642164316441645164616471648164916501651165216531654165516561657165816591660166116621663166416651666166716681669167016711672167316741675167616771678167916801681168216831684168516861687168816891690169116921693169416951696169716981699170017011702170317041705170617071708170917101711171217131714171517161717171817191720172117221723172417251726172717281729173017311732173317341735173617371738173917401741174217431744174517461747174817491750175117521753175417551756175717581759176017611762176317641765176617671768176917701771177217731774177517761777177817791780178117821783178417851786178717881789179017911792179317941795179617971798179918001801180218031804180518061807180818091810181118121813181418151816181718181819182018211822182318241825182618271828182918301831183218331834183518361837183818391840184118421843184418451846184718481849185018511852185318541855185618571858185918601861186218631864186518661867186818691870187118721873187418751876187718781879188018811882188318841885188618871888188918901891189218931894189518961897189818991900190119021903190419051906190719081909191019111912191319141915191619171918191919201921192219231924192519261927192819291930193119321933193419351936193719381939194019411942194319441945194619471948194919501951195219531954195519561957195819591960196119621963196419651966196719681969197019711972197319741975197619771978197919801981198219831984198519861987198819891990199119921993199419951996199719981999200020012002200320042005200620072008200920102011201220132014201520162017201820192020202120222023202420252026202720282029203020312032203320342035203620372038203920402041204220432044204520462047204820492050205120522053205420552056205720582059206020612062206320642065206620672068206920702071207220732074207520762077207820792080208120822083208420852086208720882089209020912092209320942095209620972098209921002101210221032104210521062107210821092110211121122113211421152116211721182119212021212122212321242125212621272128212921302131213221332134213521362137213821392140214121422143214421452146214721482149215021512152215321542155215621572158215921602161216221632164216521662167216821692170217121722173217421752176217721782179218021812182218321842185218621872188218921902191219221932194219521962197219821992200220122022203220422052206220722082209221022112212221322142215221622172218221922202221222222232224222522262227222822292230223122322233223422352236223722382239224022412242224322442245224622472248224922502251225222532254225522562257225822592260226122622263226422652266226722682269227022712272227322742275227622772278227922802281228222832284228522862287228822892290229122922293229422952296229722982299230023012302230323042305230623072308230923102311231223132314231523162317231823192320232123222323232423252326232723282329233023312332233323342335233623372338233923402341234223432344234523462347234823492350235123522353235423552356235723582359236023612362236323642365236623672368236923702371237223732374237523762377237823792380238123822383238423852386238723882389239023912392239323942395239623972398239924002401240224032404240524062407240824092410241124122413241424152416241724182419242024212422242324242425242624272428242924302431243224332434243524362437243824392440244124422443244424452446244724482449245024512452245324542455245624572458245924602461246224632464246524662467246824692470247124722473247424752476247724782479248024812482248324842485248624872488248924902491249224932494249524962497249824992500250125022503250425052506250725082509251025112512251325142515251625172518251925202521252225232524252525262527252825292530253125322533253425352536253725382539254025412542254325442545254625472548254925502551255225532554255525562557255825592560256125622563256425652566256725682569257025712572257325742575257625772578257925802581258225832584258525862587258825892590259125922593259425952596259725982599260026012602260326042605260626072608260926102611261226132614261526162617261826192620262126222623262426252626262726282629263026312632263326342635263626372638263926402641264226432644264526462647264826492650265126522653265426552656265726582659266026612662266326642665266626672668266926702671267226732674267526762677267826792680268126822683268426852686268726882689269026912692269326942695269626972698269927002701270227032704270527062707270827092710271127122713271427152716271727182719272027212722272327242725272627272728272927302731273227332734273527362737273827392740274127422743274427452746274727482749275027512752275327542755275627572758275927602761276227632764276527662767276827692770277127722773277427752776277727782779278027812782278327842785278627872788278927902791279227932794279527962797279827992800280128022803280428052806280728082809281028112812281328142815281628172818281928202821282228232824282528262827282828292830283128322833283428352836283728382839284028412842284328442845284628472848284928502851285228532854285528562857285828592860286128622863286428652866286728682869287028712872287328742875287628772878287928802881288228832884288528862887288828892890289128922893289428952896289728982899290029012902290329042905290629072908290929102911291229132914291529162917291829192920292129222923292429252926292729282929293029312932293329342935293629372938293929402941294229432944294529462947294829492950295129522953295429552956295729582959296029612962296329642965296629672968296929702971297229732974297529762977297829792980298129822983298429852986298729882989299029912992299329942995299629972998299930003001300230033004300530063007300830093010301130123013301430153016301730183019302030213022302330243025302630273028302930303031303230333034303530363037303830393040304130423043304430453046304730483049305030513052305330543055305630573058305930603061306230633064306530663067306830693070307130723073307430753076307730783079308030813082308330843085308630873088308930903091309230933094309530963097309830993100310131023103310431053106310731083109311031113112311331143115311631173118311931203121312231233124312531263127312831293130313131323133313431353136313731383139314031413142314331443145314631473148314931503151315231533154315531563157315831593160316131623163316431653166316731683169317031713172317331743175317631773178317931803181318231833184318531863187318831893190319131923193319431953196319731983199320032013202320332043205320632073208320932103211321232133214321532163217321832193220322132223223322343225322632273228322932303231323232333234323532363237323832393240324132423243324432453246324732483249325032513252325332543255325632573258325932603261326232633264326532663267326832693270327132723273327432753276327732783279328032813282328332843285328632873288328932903291329232933294329532963297329832993300330133023303330433053306330733083309331033113

FIGURE 221

GTGTGTCTCTCAGCAAAACAGTGGATTTAATCTCCTTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAACCGAACCTGACAAAAAGAAAGAAAGAAAGAA
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCCTATT
GACAACCGGGTCACCCGGGTGGCTGGCTAAACCGCAGCACCATCCTCTATGTGGGAATGA
CAAGTGGTGCTGGATCCTCGCGTGGTCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATTTT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCTTACGGTTACTTGGAGACACATCTCTCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCGTGGTACGGAGAGTAAAGGTACCCGTGAACATATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCGTGGGACAAAAGGGGACATGCAGTGTGAAGCC
TCAGCAGTCCCTCAGCAGAATTCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCTCTCAAACTCATCTTCTCAATGTCTCTG
AACATGACTATGGGAACCTACTTGGTGGCCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGTCTATTGGTCCAGGCGCCGTGAGCGAGGTGAGCAACGGCAGCTCGAGGAGGGCAGG
CTGGCTCTGGCTGCTGCTCTTCTGGTCTTGACCTGCTTCTCAAATTTGAGTGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCAACACAAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGAAAAGAGTTTTAAAAAGAAATTGAA
AATTGCTTGCAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAACGTGAAATAAAAGAGCAAAA
AAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVFVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSITILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYECASNDVAAPVVRVKVTVNYPPYISEAKGTGVPVGQKGTLCQCEASAV
PSAEFQWYKDDKRLIEGKKGVKENRPFLSKLIFFNVSEHDYGNITCVASNKLGHNTASIML
FGPGAVSEVSNGTSSRRAGCVWLLPLLVLHLLKLF

Signal peptide:

amino acids 1-28

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184

FIGURE 223

GAAAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAAGATTT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2

FIGURE 224

ATGGCTGGTGAACGGCGGGGCGGGCAGGGGACCGGGGCCGGGGAGCGGGCCAGCTGCGGGGAGCCCTGA
 ATCACCGCCTGGCCCACTCCACATGAAACGTTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACCTGGGATCCAG
 AAGGGGACAGACAGCTGTTAGGCTCACGCACGAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
 GCATCTCTCTGGCTGCTCTGTGGCCCTAGGSGTCTCAGTACCACAGAGACCCATCCACAGCACCTGCCTTACA
 GAGGCTCTGATTGAGTGGCTGGAATAATCTGAGTCCCTGGACCGAGGGGTGAGCCCTGTGAGGACTTTTAC
 CAGTTCTCTGTGGGGCTGGATTTCGAGGGAACCCCTCTGCCGATGGGCTTCTCGCTGGAACACCTTCAACAGC
 CTCTGGGACCAAAACAGGCCATCTGAAGCACCTGCTTGAAAAACACCACTTCAACTCCAGCAGTGAAGCTGAG
 CAGAAGACACAGCGCTTCTACCTATCTTGCTTACAGGTGAGCGCATTGAGGAGCTGGGAGCCAGCAGCTGAGA
 GACCTCATTTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCGAGGACACTTTATGAGAGTTTCAACAGC
 GCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTACCGTCTACATCAGTGCCGACTCTAAGAGTTTCAACAGC
 AATGTTATCCAGGTGAGACAGTCTGGGCTCTTTCTGCCCCCTCGGAGTTACTACTTAAACAGAACTGCCAATGAG
 AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACCTGGGATGCTGCTGGGTGGGCGGCCCACTTCCAGAGG
 GAGCAGATGCAGCAGGTGCTGGAGTTGAGATACAGCTGGCCAAATCAGCAGTGGCCGAGGACAGCGGCGCGAC
 GAGGAGAAGATCTACCAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTT
 CTGTCTTTCTGCTGTCAACATTTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCGAG
 CAGGTGTGAGAGCTCATCAACCGCACGGAAACCAAGCATCTGAACAATTACCTGATCTGGAACTGGTGCAAAAG
 ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCCTCTATGGCACTAAGAAG
 TCCTGTGTGCGCAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCCCTGGCTTTGCTTTGGGGTCACTCTTC
 GTGAAGGCCACGTTTGACCGGCAAGCAAGAAATTCAGAGGGGATGATCAGCGAAATCCGGACCGCATTGTAG
 GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAAGGAGAAAGCAGATGCCATCTAT
 GATATGATTTGGTTTCCAGACTTTATCCTGGAGCCCAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
 GAAGATTCTTTCTTCAAAAACATGTTGAATTTGTACAACCTCTCTGCCAAGGTTATGGCTGACCACTCCGCAAG
 CCTCCAGCCGAGACAGTGGAGATGACCCCCCAGACAGTGAATGCTACTACTCTCCAATAAGAAATGAGATC
 GTCTTCCCCGCTGGCATTCTGCAGGCCCTTCTATGCCCGCAACCAACCCAAAGGCCCTGAACTTCGGTGGCATC
 GGTGTGGTATGGGCCATGAGTTGACGCAATGCTTTGATGACCAAGGGCGCGAGTATGACAAAGAGGGAACTTG
 CGCCCTGGTGGCAGAAATGAGTCCCTGGCAGCTTTCGGAACCAACGGCTGATGGAGGAACAGTACAATCAA
 TACCAAGTCAATGGGGAGAGGCTCAACGGCCGCAAGACGCTGGGGGAGAACATTACTGACAAACGGGGGCTGAAG
 GCTGCCCTAATGCTTACAAGCATGGCTGAGAAAGCATGGGGAGGACAGCACTCTCCAACCTCCGTGACTCTCTGGG
 AACCCACAGCTCTTCTCTGCTGGGATTTGCCAGGTGTGGTCTCGGTCCGCAACAGAGAGCTCTCAGAGGGG
 CTGGTGACCGAGCCCAAGCCCTGCCGCTTCCGCGTGTGGGCACTCTCCAACCTCCGTGACTCTCTCTGGG
 CACTTCGGCTGCCCTGTCCGCTCCCCATGAACCCAGGCGAGCTGTGTGAGGTGTGGTGGAGCTGGATCAGGGGA
 GAAATGGCCAGCTGTCAACAGCTCTGGGGCAGCTCTCTGACAAAGCTGTTTGTCTTGGGTGGGAGGAAGCAA
 ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCAGAGGTGACATGAGTACAGCCCTCTCAATCAACCAATTG
 TGCCTCTGCTTTGGGGTGCCTGCTCCAGCAGAGCCCCCACTTCACTGTGACATCTTTCCGTGTCAACCT
 GCTTCGAAGAGGTCTGGGTGGGGAGCCAGTTCCCATAGGAAGGAGTCTGCG

FIGURE 225

MNVALQELGAGSNVGFQKGTROLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDP SH
STCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGGWIRRNPLDGRSRWNTFNSLWDQNQA
ILKHLENTTFNSSEAEQKTQRFYLSCLQVERIEELGAQPLRD LIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFTVYIISADSKSSNSNVIQVDQSG LFLPSRDYYLNRTANEKVLTAY
LDYMEELGMLLGGRPTSTREOMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPELESDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLETLTYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLNLYNSAKVMADQLRKPPSRDQWSMTPTQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLEFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRLVGLTLSNSRDFLRHFGCPVGS PMNPGQLCEVW

Type II Transmembrane domain:

amino acids 32-57

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCGGTGATGGCTGGTGACGCGCGGGCCGGGCAGGGGA
CCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCACAGCACCTGCCTTACAGAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCGTGTGAGGACTTTTACCAGTTCTCCTGTGGGGCTGGATTTCGGA
GGAACCCCTGCCCAGTGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAAACCACTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCATTCTTAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCACACGAATGTTATCCAGGTGGACCACTG
GGCTCTTCTGCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGTCAGAGCAGGAAGGTGAGCCTATCCTGTACCTAGTGAACAAACTGCCCT
CCTTCTTTCTTCTTTCTTCTCCTCCCTCCCTCCCTTTCTTCCCCTTTTCTTCTTCTTCTTCC
TCTTATCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTGC
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTGATGGGTTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCTGAGGATTGAGGAGAGCTTATGGAGCC
AGCAAAGTCTTCTGAAGAGATTGCATTTGAGCCAGGTCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAATAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACCTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGCGCTGCATGGAGGAACAG
TACAATCAATAACAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCAACTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTGCGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACACATTGTGCCTCTGCTTTGGGGGTGCCCT
GCCTCCAGCAGAGCCCCACCATTCACTGTGACATCTTTCGTGTACCCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTGTGCCTCTTCTGTCCCAGGCTCACT
CAGCCTGGCGCCATGGGGCTGCCGTGCCTGCCCACTGTGACCCACAGGCCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCACACTTAGGGGTGGA CTGACTCAGCTCTGTC
TGECTCACCTCACGGGCTACCCCCACCTCACCTGTGCTCCTTGTGCCACTGCTCCAGTGT
CTGCTGTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTTGAAAGCCTCCTGC
TGCCCACTGTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGTACTGGTTCCGTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTACAGAAAAGAGGTGGGAGGGTGTGCTCT
TGGCCCTTATAGGACC

FIGURE 229

CCACGCGTCCGAGCCGCCGAGAAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACACCGAAAAACAAAAGAGAGAAACACACCCAACACTGGGGGTGG
GGGGAAGAAAGAAAGAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCTGTGGCGCGCCCTGGTTCCCGGGAAGACTCGCCAGCACCGGGGG
TGGGGGAGTGCAGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAAGGGTGCTTGTGTGCTCGAACAGTGGCTGGCGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCTCCTGCTCCCGGCTGGACAGAGTGTGGACTTCCCTGGGGCGCGGTGG
ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCTGGCTGAACCGGTCAAGTATTATTTTGGCGAGGTGATGAAGTGTCAAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACAGTGTCTGTTCAGACTCAACATACACCCAGAACA
ATGCAAGTGTCATCTAACTGTGCAAGTTCCCTCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACCTTACTTGTTTGGCCACTGGGAAACAGAGCCTTCCA
TTTCTTGGCGACATCTCCCCATCAGCAAAACCATTTGAAAAATGGACAATATTTGGACATT
TATGGAATTTACAGGGACACAGGCTGGGGAAATGAATGCACTGCGGAAAAATGCTGTGTCAAT
CCCAGATGTGAGGAAAGTAAAAGTTGTGTCAACTTGTCTTACTATTTCAGGAAATTAAT
CTGGCACCGTGACCCCGGACGCAAGTGGCTGATAAGATGTGAAGGTGCAAGGTGTGCCGCC
CCAGCCTTTGAATGGTACAAGAGGAGAGAAGCTCTTCAATGGCCAAACAGGAATTATTAT
TCAAAATTTTACCAACAGATGTCATTCTCACTGTTACCAACGTGACACAGGAGCATCTCGGCA
ATTATACCTGTGTGGCTGTGCCAACAGCTTAGGCAACAACATGCGAGCCTGCCCTCTTAACCT
CCAAGTACAGCCAGTATGGAATTAACGGGAGCGCTGATGTTCTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCTCTTTTCCACAGCATATTTCTACCTGAAGAAATGCCATTCTACAATAA
TTCAAAGACCCATAAAAGGCTTTTAAGGATTTCTGAAAAGCTGTGATGGCTGGATCCAATCT
GGTACAGTTTGTAAAAGCAGCGTGGGATATAATCAGCAGTGCCTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAATACTTTAATTCTACTCTTTTGGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCTCTTTTAAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTTGTGATTTTCAATGTTTGAATCTACAACCTTTTCAAAGCATTTCAGTCATGGT
CTGCTAGGTTGCAAGCTGTAGTTTACAAAAACGAATATTGCAAGTGAATATGTGATTTCTTAA
GGCTGCAATACAAGCATTCAAGTTCCTCTGTTTCAATAAGAGTCAATCCACATTTACAAGATG
CATTTTTTCTTTTTTGATAAAAAAGCAATAATATTGCCTTCAGATTAATTTCTTCAAATA
TAACACATATCTAGATTTTCTGCTTGCTGATGATATTCAAGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTCTCTCTTCTGTAAGTTTCAGCATGGGTGTGCCCTCATAC
AATAATATTTTCTCTTTGTCTTCAACTAATAAAATGTTTGTCTAAATCTTACAATTTGA
AAGTAAAAATAAACACAGAGTGATCAAGTTAAACCATACACTATCTCTAAGTAAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCTCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT
AACTCACTTCTGTGTGATGAGACAAATTTAATAACAGTATAGTAAATATACCATATGATTTT
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCTTTTAAATGACAGCACA
GTCCACTCAAAGGATTTGCCTAGCAATACAGCACTCTTTCTTCTACTAGTCCAAGCCAAAA
TTTTAAGATGATTTGTGAGAAAGGGCAAAAGTCTTACCTAATATTACAAGAGTTGGTA
AGCGCTCATCAATTAATTTATTTGTGGCAGGTATTATGACAGTGCAGCTGGAGGGTATGGA
TATGGATATGGACATTTTCCAGAGCTATAATGGCAGAAACAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACCTGAAATGAGACATGCACATAAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCTCTCCAAATGGCTGTATTTATAAAGGTTTGTG
AGCTGCACTGAAGACTCTTATTTTATAGTATATCAACCTTTTGTTTTAAATTTGACCTGGCA
AGGTAGCTGAAGACCTTTTAGACAGTTCATCTTTTCTTAAATTTTTTCTGCTTATTTAA
AGACAAATTTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLA AVL LSLCCLLPAGQSVDFPWA AVDNMMVRKGDTAVLR CYLED
GASKGAWLNRSSII FAGGDKVSDPRVSISTLNKRDYSLQIQNV DVTDDGPYTC SVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFDPVRKVKVVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQEHFGNYTCVAANKLGT TNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLS SFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTCTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAAAGCTGCTTTTACCACAAATTGCAATGGAGCCTTTTCGAAA
TCAATGTTCCAAAGCCCCAAGAGGAGAAATGGGGTGAACCTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCTTGAGATGTATTTCCTCAATGACACTCTGGCGGTGAGGACAGCCGCT
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACCTTCACTCAGAACCAGGGATGTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAAGTCAACAAGGGGGCCATGGGCATGCCTGGTGCCCCCTGGCCCGCCGGGACCACTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCTCGGGACCCCAAGG
CCCACCGGAGTCAAGGGAGAGGCGGGCTTCCAAGGACCCAGGGTGTCCAGGGAAGCAAG
GAGCCACTGGCACCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCCAAAGGGGAACTGGAACTAAGGGAGAGAAGGAGACCTGGGTCTCCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTATGGGGCTCCTGGAGCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACCAGGTTTGGCTGGTCTTCTTGAGCTAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCGGGGCCCTCCTGGTGCACTGGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCTGGGCGAGCAGGACTTCAGGGAGCCCCCGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGAACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCAGG
GCTGCGAGGTCCCAAGGGAGCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAGAGGTGAAACTCAGTGTCC
GTCAGGATTGTGCGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATTTCTGATGCCATTGTCTTCTGCCGATGCTGGGTT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTCAGTGTGCGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCGAGGCGTGGAGTGCAGCGCTCTGACCCGGAAACCCTTTCA
CTTCTCTGCTCCCGAGGTGTCTCCTGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAPHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLLITAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAAGAMGRD
GATGPGSPGQPPGPKGEAGLQGPQGAAPGKQGATGTPGPQGEKSGKGDGGLIGPKGETGTKGE
KGDGLPGSKGDRGMKGDAVVMGPPGAQGSKGDGFRPGPPGLAGFPAGKGDQGPGLQGVPG
PPGAVGHGPAKGEFGSGAGSPGRAGLPGSPGSPGATGLKSGKGDGTLQGGQGRKGESGVPGPA
GVKGEQGSPPGLAGPKGAPGQAGKGDQGVKGSSGEGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLNDNVQCRGTESTLW
SCTKNSWGHHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCAATTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGCCAAACTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAGGAAAGAAGAAAAAGGCCAAAAGCCAAAATGAACTGATGGTACTTGTTTTCAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTCAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCAAAAGACGTTTTCTTTGGACCAAAGATCTCTTTG
TGATTCTTGCAACAATCAATGAATCTTCATGTATTCTGGAGAACACCATTCTGATTTC
CCACAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAA

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758

<subunit 1 of 1, 98 aa, 1 stop

<MW: 11081, pI: 6.68, NX(S/T): 1

MKLMVLVFTIGLTLLLGTVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG

KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGCTCTGGAGCGAATTCACAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCGTGGTGGTTGGAGGGC
GCGCAGTAGAGCAGCAGCAGCGCGCGGGTCCCGGGAGGCGCGCTCTGCTCGCGCCGAGATG
TGGAAATCTCTTACAGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTCTGGCGGGTGGCTTCTTCTCCTCGGCTTCTCTTGGGTGGTTTA
TAAATCTCCTCAATGAAGCTACTAACATTACTCCAAGCATATAATGAAAGCATTTTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCAATTTAGC
AGGAACAGAACAAATCTTTCAGCTTGCAAAGCAAATTCAATCCAGTGGAAGAATTTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCTCTGTTGTCTTACCCAAATAAGACTCATCCC
AATCATACTCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTAATTTGAACC
ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTTCTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGATGTAACTATGCACGAACCTGAAGACTTCTTTAAA
TTGGAACGGGCATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
CCGACCTTGCTGACTACTTTGCTCTCGGGTGAAGTCCCTTCAGAGCGTTGGAATCTTCTCT
GGAGGTGTTGTCAGCGTGGAAATATCCTAAATCTGAATGGTGAGAGACCTCTCACACC
AGGTTACCGCAGCAATGAATATGCTTATAGGCGTGAATTTGCAGAGGCTGTTGGTCTTCCAA
GTATTTCTGTTTCATCCAATTGGATACTATGATGCACAGAAGCTCTAGAAAAAATGGGTGGC
TCAGCACCACAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
CTTTACTGGAAACTTTTCTACAAAAAGTCAAGATGCACATCCACTCAACATGAAGTGA
CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGAACCAGACAGATATGTCAAT
CTGGGAGGTACCGGGGACTCATGGGTGTTTGGTGGTATTGACCTCAGAGTGGAGCAGCTGT
TGTTTCATGAAATTTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTTTGTGTTGCAAGCTGGGATGCGAAGAATTTGGTCTTCTTGGTCTACTGAGTGGGCA
GAGGAGAATTCAGACTCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
AGAAGGAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAATCTCTTTATGAAAGTTGG
ACTAAAAAAGTCCCTCCCGAGGTTTCAAGTGGCATGCCAGGATAAGCAAAATTTGGGATCTGG
AAATGATTTTGAGGTGTTCTTCCAACGACTTGAATGTCTTCAGGCAGAGCACGGTATACTA
AAAATTTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCAGTGTGGCCAGGTTTCG
AGGAGGGATGGTGTGTTGAGCTAGCCAATTCATAGTGCTCCCTTTGATTGTGCGAGATTATG
CTGTAGTTTAAAGAAATGCTGACAAAAATCTACAGATTTCATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCATTTGATTCACTTTTTCTGTCAGTAAAGAATTTTACAGAAAT
TGCTTCCAAGTTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAAATAGTATTAAGAA
TGATGAATGATCAATCATGTTTCTGGAAGAGCATTATTGATCCATTAGGGTTACAGAGC
AGGCCTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
ATTCACAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGGACCTTCCAAGGCCCT
GGGAGAGAGTGAAGAGACAGATTATGTTGCAAGCTTTCAGAGTCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCCTTAAGAGGATTTTATAGAGAAATCCGTAATTTGAGTGTGATGTCA
CTCAGAAAGAACTCGTAATGGGTATATTGATAAATTTTAAATTTGGTATATTTGAAATAAAGT
TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLETDSAVATARRPRWLGCAGLVLAGGFLLGFLFGWFIKSSNEATNITPKHNMKAFI
DELKAENIKKFLHNFTQIPLHAGTEQNFQLAKIQSQWKEFGLDSELAHYDVLVSYNPKTH
PNYISIINEDGNEIFNTSLFEPPEPPGYENVSDIVPPSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSCGKIVARYGKVFGRGNKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANAYARRGIAEAVGLPSIPVHPIGYYDAQKLEKMG
GSAPPDSSWRGSLKVPYNVGPFGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDYV
ILGHRDSWVFGGIDPQSGAAVVEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKEKSPDEGFEGKSLYES
WTKKSPSPFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRYADKIYISIMKHPQ
EMKTYSVSFDLSFSAVKNFTEIASKFSERLQDFDKSNPIVLRMMDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHKNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713